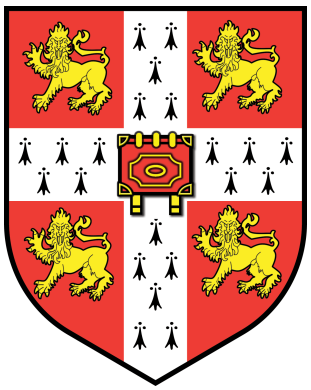
Numerical Solution of Sturm–Liouville Problems via Fer Streamers

Alberto Gil Couto Pimentel Ramos

Homerton College

University of Cambridge



A thesis submitted for the degree of

Doctor of Philosophy

April 2016

Abstract

The subject matter of this dissertation is the design, analysis and practical implementation of a new numerical method to approximate the eigenvalues and eigenfunctions of regular Sturm–Liouville problems, given in Liouville's normal form, defined on compact intervals, with self-adjoint separated boundary conditions.

These are classical problems in computational mathematics which lie on the interface between numerical analysis and spectral theory, with important applications in physics and chemistry, not least in the approximation of energy levels and wave functions of quantum systems.

Because of their great importance, many numerical algorithms have been proposed over the years which span a vast and diverse repertoire of techniques. When compared with previous approaches, the principal advantage of the numerical method proposed in this dissertation is that it is accompanied by error bounds which:

(i) hold uniformly over the entire eigenvalue range, and,

(ii) can attain arbitrary high-order.

This dissertation is composed of two parts, aggregated according to the regularity of the potential function. First, in the main part of this thesis, this work considers the truncation, discretization, practical implementation and MATLAB software, of the new approach for the classical setting with continuous and piecewise analytic potentials (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*). Later, towards the end, this work touches upon an extension of the new ideas that enabled the truncation of the new approach, but instead for the general setting with absolutely integrable potentials (Ramos, 2014).

Keywords

Numerical method; Eigenvalues; Eigenfunctions; Regular Sturm–Liouville problems; Liouville's normal form; Self-adjoint separated boundary conditions; Continuous and piecewise analytic potentials; Absolutely integrable potentials; Uniform; High-order; Fer expansions; Fer streamers; Lie-algebraic techniques; Multivariate oscillatory quadrature; Reduced Hall basis; MATLAB;

Mathematics subject classification

65L15; 65L10; 65L70; 34C40;

I dedicate this thesis to my mother Sofia and to my father Armando.

Acknowledgements

First and foremost, I would like to thank my PhD advisor Prof. Arieh Iserles. His unwavering love for mathematics and everlasting devotion to research, have been an immense source of inspiration throughout my PhD research. I am especially thankful for his encouragement and support, which directed me towards the areas of Numerical Analysis and Spectral Theory, and to the Sturm–Liouville problems that form the subject of this dissertation.

My gratitude and appreciation go also to the members of the Cambridge Numerical Analysis group, at the Department of Applied Mathematics and Theoretical Physics, at the University of Cambridge, with whom I have had the pleasure to interact with. Special thanks go to my scientific brothers, Pranav Singh and Marcus Webb, as well as to Dr. Andreas Asheim and Dr. Karolina Kropielnicka, for useful discussions.

I have also been fortunate to be a part of the Cambridge Centre for Analysis (CCA), at the Department of Applied Mathematics and Theoretical Physics, at the University of Cambridge. Most importantly, I am indebted to my friends at CCA, with whom I have shared almost every day of the last four years. They have given me the emotional support and fortitude that have propelled me throughout my PhD life. First, I thank Milana Gatarić for being the best friend one could hope for. I have benefited much from having her as my office companion and part of my daily life, and I remain always inspired by her commitment to research. Beside Milana, I thank also: Luca Calatroni, Alberto Coca, Kevin Crooks, Eoin Devane, Clarice Poon, Vittoria Silvestri and Lukáš Vermach. I have found in them friendships that will remain long after our journey through Cambridge.

Prior to my time in Cambridge, I would also like to acknowledge my Master advisor Dr. Miguel Raul Dias Rodrigues, for believing in me and introducing me to research during our time at the University of Porto.

Most importantly, I thank my mother Sofia, father Armando, sister Sara and brother Filipe for their encouragement to pursue my dreams. Without their constant support, I would have never reached the University of Cambridge.

Finally, I would like to acknowledge also the support of the Fundação para a Ciência e a Tecnologia, Portugal, through the fellowship SFRH/BD/71692/2010.

Alberto Gil Couto Pimentel Ramos
Cambridge
14 December 2015

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution.

Contents

List of notation

Italic lower case letters denote scalars, boldface lower case letters denote column vectors and boldface upper case letters denote matrices, e.g., c, $\boldsymbol{c}$ and $\boldsymbol{C}$, respectively. The transpose of $\boldsymbol{C}$ is denoted by $\boldsymbol{C}^\top$. Given two matrices, with the same dimensions, $\boldsymbol{M}_1, \boldsymbol{M}_2 \in \mathbb{C}^{j_1 \times j_2}$, the Hadamard product between $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ is denoted by $\odot$ and is the unique element in $\mathbb{C}^{j_1 \times j_2}$ defined by $[\boldsymbol{M}_1 \odot \boldsymbol{M}_2]_{j_3,j_4} := [\boldsymbol{M}_1]_{j_3,j_4} [\boldsymbol{M}_2]_{j_3,j_4}$.

$[a,b]$, q, $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$	Interval, potential and boundary conditions of the regular Sturm–Liouville problem (1.0.1)–(1.0.2), Page 1
(λ, y_λ)	Generic eigenvalue and eigenfunction pair of the regular Sturm–Liouville problem (1.0.1)–(1.0.2), Page 1
$(\lambda_j, y_{\lambda_j})$	j-th eigenvalue and eigenfunction pair of the regular Sturm–Liouville problem (1.0.1)–(1.0.2), Page 1
$q_{\min}$, $q_{\max}$	Lower bound to the minimum of q, and upper bound to the maximum of q, for continuous and piecewise analytic q, Page 6
$z_{\lambda,h}$, $\varpi_{\lambda,h}$	Eq. (1.1.4)
$\mathrm{SL}(2,\mathbb{R})$	Matrix representation of the Lie group of two-by-two real matrices with determinant one, Eq. (1.1.9)
$\mathfrak{sl}(2,\mathbb{R})$	Matrix representation of the Lie algebra of two-by-two real matrices with zero trace, Eq. (1.1.10)

$\rho(\boldsymbol{X})$, $\exp(\boldsymbol{X})$, $\mathrm{Ad}_{\exp(\boldsymbol{X})}\boldsymbol{Y}$, $\mathrm{ad}_{\boldsymbol{X}}\boldsymbol{Y}$, $[\boldsymbol{X},\boldsymbol{Y}]$	Definition 2.1.1
$\boldsymbol{B}_{\lambda,l}(c_k,t)$, $\boldsymbol{D}_{\lambda,l}(c_k,t)$	Definition 2.1.2
$\boldsymbol{\pi}(\boldsymbol{X})$, $\boldsymbol{\pi}^{-1}(\boldsymbol{x})$, $\mathscr{C}_{\boldsymbol{X}}$	Definition 2.1.3
$\psi(z)$, $\varphi(z)$, $\phi(z)$	Definition 2.1.4
$\delta_{\lvert q'\rvert}$	Definition 2.1.5
$\boldsymbol{F}_{\lambda}(c_k,c_{k+1})$	Definition 2.1.6
$\boldsymbol{Y}_{\lambda}(c_{k+1})$	Solution to the initial value problem (1.0.4)–(1.0.5) evaluated at $t = c_{k+1}$, Definition 2.1.6
$\tilde{\boldsymbol{F}}_{\lambda,n}(c_k,c_{k+1})$, $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$	Definition 2.1.6
$\boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k,c_{k+1})$, $\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_{k+1})$	Definition 2.1.6
$\zeta_{\lambda,1}(c_k,t)$, $\boldsymbol{R}_1$, $\boldsymbol{S}_{\lambda,1}(c_k,t)$, $\boldsymbol{U}_{\lambda,1}(c_k,t)$, $\boldsymbol{V}_{\lambda,1}(c_k,t)$	Definition 3.2.1
$\omega_{\lambda,1}(c_k,t)$, $r_{\lambda,1}(c_k,t)$, $\epsilon_{\lambda,1}(c_k,t)$, $s_{\lambda,1}(c_k,t)$	Definition 3.3.1
$\boldsymbol{f}_{\lambda,1}(c_k,t)$, $\boldsymbol{\iota}_{\lambda,1}(c_k,t)$, $\boldsymbol{g}_{\lambda,1}(c_k,t)$	Definition 3.3.2
$\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k,c_k+h_kt)$	Definition 3.3.3
$\tilde{\boldsymbol{D}}_{\lambda,j,n}(c_k,c_{k+1})$, $\boldsymbol{E}_{\lambda,j,n}(c_k,c_{k+1})$	Definition 3.4.1
$\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k,c_{k+1})$, $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$	Definition 3.4.2, Definition 4.1.3
$\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})$, $\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_{k+1})$	Definition 3.4.2
$\boldsymbol{L}^{\text{total}}_{\lambda,n}(c_k,c_{k+1})$, $\boldsymbol{G}^{\text{total}}_{\lambda,n}(c_{k+1})$	Definition 3.4.3
$\mathcal{S}_{l-2}$, $\mathcal{T}_{l-1}$, $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k,c_k+h_kt)$	Definition 4.1.1
$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k,c_{k+1})$, $\tilde{\boldsymbol{I}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k,c_{k+1})$	Definition 4.1.2
ϵ_1, ϵ_2	Definition 6.3.1

Chapter 1

Introduction

This dissertation is about a new numerical algorithm, named Fer streamers, to approximate the eigensystem of regular Sturm–Liouville problems, in Liouville's normal form, defined on compact intervals

$$\begin{gathered} -y_\lambda''(t) + q(t)y_\lambda(t) = \lambda y_\lambda(t), \quad t \in [a,b], \quad a,b \in \mathbb{R}, \quad \lambda \in \mathbb{R}, \\ q : [a,b] \to \mathbb{R}, \quad y_\lambda : [a,b] \to \mathbb{R}, \end{gathered} \tag{1.0.1}$$

with self-adjoint separated boundary conditions

$$\begin{aligned} \alpha_1 y_\lambda(a) + \alpha_2 y_\lambda'(a) = 0, \quad \alpha_1, \alpha_2 \in \mathbb{R}, \quad \alpha_1^2 + \alpha_2^2 > 0, \\ \beta_1 y_\lambda(b) + \beta_2 y_\lambda'(b) = 0, \quad \beta_1, \beta_2 \in \mathbb{R}, \quad \beta_1^2 + \beta_2^2 > 0, \end{aligned} \tag{1.0.2}$$

where a, b, q, α_1, α_2, β_1 and β_2 are known, and the challenge is to approximate numerically the unknown eigenvalue and eigenfunction pairs (λ, y_λ) (Pryce, 1993; Zettl, 2005).

It has been known for many years that if q is absolutely integrable then (1.0.1)–(1.0.2) possess a unique countable family of solutions

$$\left\{ (\lambda_j, y_{\lambda_j}) : j \in \mathbb{Z}_0^+,\ \lambda_j \le \lambda_{j+1} \text{ and } \|y_{\lambda_j}\|_{\mathrm{L}^2([a,b],\mathbb{R})} = 1 \right\}.$$

Because of this, there are in fact two different numerical approximation problems associated with (1.0.1)–(1.0.2): Given $\epsilon > 0$ and

(a) given a compact interval $[\lambda_{\min}, \lambda_{\max}] \subseteq \mathbb{R}$, compute the eigenvalues λ_j of (1.0.1)–(1.0.2) in this interval approximately with ϵ precision together with their corresponding eigenfunctions y_{λ_j} also with ϵ precision (pointwise or in $\mathrm{L}^2([a,b],\mathbb{R})$), or alternatively

(b) given two indices $j_{\min}, j_{\max} \in \mathbb{Z}_0^+$ with $j_{\min} \le j_{\max}$, approximate with ϵ precision the eigenvalues λ_j of (1.0.1)–(1.0.2) with j between $j_{\min}$ and $j_{\max}$ together with their corresponding eigenfunctions y_{λ_j} also with precision ϵ.

These are classical problems in computational mathematics, ubiquitous in applications, important in physics, chemistry and applied mathematics, e.g., in fluid flow, Schrödinger spectra, nuclear magnetic resonance imaging, etc (Amrein, Hinz and Pearson, 2005).

In both problems, i.e., either (a) given the task to approximate all eigenvalues with values between $\lambda_{\min}$ and $\lambda_{\max}$ or alternatively (b) given the task to approximate the eigenvalues with indices between $j_{\min}$ and $j_{\max}$, the approach is to capitalize on representations of the eigenvalues as the roots of certain equations and to call upon root-finding techniques to compute the eigenvalues.

Most likely, the reader is at this moment inquiring about two very important things. Firstly, which equations are there to root-find, and are they available exactly or do they require approximation? Secondly, what information is there about the multiplicity of the eigenvalues when viewed as the roots of such equations?

An answer to the second question follows from the fact that the eigenvalues satisfy (Pryce, 1993; Zettl, 2005, Theorem 4.6.2)

$$\lambda_j < \lambda_{j+1} \text{ and } \lim_{j\to+\infty} \lambda_j = +\infty,$$

i.e., they are simple, bounded from below and accumulate only at infinity. In particular, the roots are simple, and root-finding needs not worry about multiple roots[1].

As for an answer to the first question, starting with which equations characterize the eigenvalues as their roots, there are two types:

(i) those that do not 'count' the number of oscillations in the eigenfunctions, which can be employed to tackle problem (a) above, and,

(ii) those that do 'count' the number of oscillations in the eigenfunctions, which are necessary to tackle problem (b) above.

Both types are based on writing the differential equation in (1.0.1) in the system form

$$\begin{bmatrix} y_\lambda(t) \\ y'_\lambda(t) \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ q(t)-\lambda & 0 \end{bmatrix} \begin{bmatrix} y_\lambda(t) \\ y'_\lambda(t) \end{bmatrix} \tag{1.0.3}$$

and considering either:

(i) the fundamental matrix associated with (1.0.3) or

(ii) a representation of (1.0.3) in certain polar coordinates known as 'Prüfer variables'.

[1]The root-find is in fact not so trivial, since there exist problems where the roots — although simple — can be closer than machine precision, or any other fixed small constant. For example, it is known that certain parameter dependent problems can suffer from such issues. This is the case for instance with the eigenvalues of the Coffey–Evans problem that appears in Sections 3.5 and 5.5, if its parameter β is taken large enough as happens in applications. There are ways however of dealing with such clustering issues, which have been implemented in the MATLAB package that comes with this dissertation, c.f., Chapter 5.

Based on the fundamental matrix associated with (1.0.3), the initial value problem

$$\boldsymbol{Y}'_\lambda(t) = \begin{bmatrix} 0 & 1 \\ q(t)-\lambda & 0 \end{bmatrix} \boldsymbol{Y}_\lambda(t), \quad t \in [a,b], \quad a,b \in \mathbb{R}, \quad \lambda \in \mathbb{R},$$
$$q : [a,b] \to \mathbb{R}, \quad \boldsymbol{Y}_\lambda : [a,b] \to \mathbb{R}^{2\times 2}, \tag{1.0.4}$$

with the initial condition

$$\boldsymbol{Y}_\lambda(a) := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{1.0.5}$$

yields the classical eigenvalue characterization (1.0.6) presented in the following theorem:

Theorem 1.0.1 (Zettl, 2005, Lemmas 3.2.1–3.2.2)**.** *The unknown λ_j are given by*

$$\{\lambda_j\}_{j\in\mathbb{Z}_0^+} = \{\lambda \in \mathbb{R} : \eta_\lambda = 0\}, \tag{1.0.6}$$

where $\lambda \mapsto \eta_\lambda$ is the entire function defined by

$$\eta_\lambda := \det\left(\begin{bmatrix} \alpha_1 & \alpha_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \beta_1 & \beta_2 \end{bmatrix} \boldsymbol{Y}_\lambda(b) \right).$$

Alternatively, via a representation of (1.0.3) in certain polar coordinates, one can have a different representation of the unknown $\{\lambda_j\}_{j\in\mathbb{Z}_0^+}$, which makes use of the fact that (Zettl, 2005, Theorem 4.6.2) the eigenfunction corresponding to the j-th eigenvalue

$$y_{\lambda_j} \text{ has exactly } j \text{ zeros in } (a,b).$$

As is well-known in the literature, the idea is ingenious: in its simplest form, an unscaled Prüfer transformation represents $(y'_\lambda(t), y_\lambda(t))$ in the polar coordinates:

$$y'_\lambda(t) =: r_\lambda(t)\cos(\theta_\lambda(t))$$
$$y_\lambda(t) =: r_\lambda(t)\sin(\theta_\lambda(t))$$

which in turn recasts (1.0.3) into the system

$$\begin{aligned} \theta'_\lambda(t) &= \cos^2(\theta_\lambda(t)) + (\lambda - q(t))\sin^2(\theta_\lambda(t)) \\ r'_\lambda(t) &= (1+q(t)-\lambda)\sin(\theta_\lambda(t))\cos(\theta_\lambda(t))r_\lambda(t) \end{aligned} \tag{1.0.7}$$

and yields the seminal eigenvalue characterization (1.0.8) given in the following theorem:

Theorem 1.0.2 (Zettl, 2005, Theorems 4.5.3 and 4.6.2)**.** *Set* $\alpha \in [0, \pi)$ *and* $\beta \in (0, \pi]$ *as*

$$\begin{aligned} \tan(\alpha) &:= -\alpha_2/\alpha_1 \quad \text{if} \quad \alpha_1 \neq 0, \quad \text{and} \quad \alpha := \pi/2 \quad \text{if} \quad \alpha_1 = 0, \\ \tan(\beta) &:= -\beta_2/\beta_1 \quad \text{if} \quad \beta_1 \neq 0, \quad \text{and} \quad \beta := \pi/2 \quad \text{if} \quad \beta_1 = 0. \end{aligned}$$

Then each unknown eigenvalue λ_j, $j \in \mathbb{Z}_0^+$, *is the unique solution* $\lambda = \lambda_j$ *of the equation*

$$\theta_\lambda(b) = \beta + j\pi, \tag{1.0.8}$$

where, for each $\lambda \in \mathbb{R}$, $t \mapsto \theta_\lambda(t)$ *is the solution of* (1.0.7) *determined by the initial condition*

$$\theta_\lambda(a) := \alpha.$$

Furthermore, $\theta_\lambda(b)$ *is continuous and strictly increasing in* λ.

With Theorems 1.0.1–1.0.2 in hand, one has representations of the eigenvalues as the roots of either (1.0.6) or (1.0.8). The first represents the eigenvalues as the roots of the oscillatory function $\lambda \mapsto \eta_\lambda$, whereas the second yields the strictly increasing function $\lambda \mapsto \theta_\lambda(b)$ which gives the j-th eigenvalue as its pre-image of $\beta + j\pi$.

However, this is not enough since $\lambda \mapsto \eta_\lambda$ in (1.0.6) requires $\lambda \mapsto \boldsymbol{Y}_\lambda(b)$ exactly and similarly (1.0.8) needs $\lambda \mapsto \theta_\lambda(b)$ exactly. These unfortunately are not readily available and instead require approximation.

Without going into too many details, not to clutter the main ideas, it turns out that to approximate $\lambda \mapsto \boldsymbol{Y}_\lambda(b)$ in (1.0.6) and $\lambda \mapsto \theta_\lambda(b)$ in (1.0.8), it is sufficient to approximate instead

$$(\lambda, t) \mapsto \boldsymbol{Y}_\lambda(t), \tag{1.0.9}$$

for all $(\lambda, t) \in \mathbb{R} \times [a, b]$.

Indeed, an approximation to (1.0.9) clearly suffices to yield an approximation to $\lambda \mapsto \boldsymbol{Y}_\lambda(b)$. Similarly, to approximate $\lambda \mapsto \theta_\lambda(b)$ it is known that it is enough to approximate $(\lambda, t) \mapsto (y'_\lambda(t), y_\lambda(t))$ (Pruess and Fulton, 1993, p. 364–367; Ixaru, De Meyer and Berghe, 1999, p. 263–265), which itself can be derived easily from an approximation to (1.0.9), together with

$$\begin{bmatrix} \alpha_1 & \alpha_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_\lambda(a) \\ y'_\lambda(a) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} y_\lambda(b) \\ y'_\lambda(b) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{1.0.10}$$

$$\begin{bmatrix} y_\lambda(t) \\ y'_\lambda(t) \end{bmatrix} = \boldsymbol{Y}_\lambda(t) \begin{bmatrix} y_\lambda(a) \\ y'_\lambda(a) \end{bmatrix}, \tag{1.0.11}$$

and a suitable normalization.

Because of this, in order to approximate the eigenvalues and eigenfunctions of the Sturm–Liouville problems (1.0.1)–(1.0.2), the focus of this dissertation is to develop high-order approximations to (1.0.9) that hold uniformly in (λ, t).

With the uniform and high-order approximations to (1.0.9) constructed in this thesis, one can then approximate uniformly and efficiently both $\lambda \mapsto \boldsymbol{Y}_\lambda(b)$ and $\lambda \mapsto \eta_\lambda$ in Theorem 1.0.1 as well as $\lambda \mapsto \theta_\lambda(b)$ in Theorem 1.0.2.

This in turn permits to root-find an approximation of either (1.0.6) or (1.0.8), and compute the eigenvalues via value or index, i.e., by (a) or (b) in page 1. Finally, having approximated the eigenvalues, one can then estimate the eigenfunctions via (1.0.10)–(1.0.11).

With the outline from the three previous paragraphs in mind, it is important to distinguish between the focus of this thesis which is to approximate (1.0.9) uniformly in $(\lambda, t) \in \mathbb{R} \times [a, b]$, versus the much simpler problem to approximate

$$t \mapsto \boldsymbol{Y}_\lambda(t), \tag{1.0.12}$$

for a particular fixed $\lambda \in \mathbb{R}$ and all $t \in [a, b]$.

At least in principle, for a particular fixed λ, any ordinary differential equation solver can be called to approximate (1.0.12) at any $t \in [a, b]$. This will certainly be highly inefficient (or perhaps even impractical) for almost any solver if $|\lambda|$ is large (since (1.0.12) becomes either exponentially large or extremely oscillatory), but is nonetheless conceivable, in the sense that if one partitions the interval $[a, b]$ into tiny subintervals of size h (which will depend on the fixed λ) then error control can be guaranteed (since for each fixed λ, the step size h will be chosen small enough, as a function of λ, to yield prescribed accuracy).

Regardless of efficiency issues (which also need to be addressed), a more delicate problem is that one might not know what λ is, and therefore might not have the necessary information to restrict h in view of λ, to proceed as summarized in the paragraph above.

With this in mind, the reason for approximating the more intricate function (1.0.9) rather than (1.0.12), becomes clear when distinguishing between the two approximations required in Sturm–Liouville problems (1.0.1)–(1.0.2), i.e., (a) and (b) in page 1.

Again disregarding efficiency issues, problem (a) in page 1 is not necessarily dependent on having an approximation to (1.0.9) that yields uniform error bounds independent of λ, since the interval $[\lambda_{\min}, \lambda_{\max}]$ that is provided as part of the problem formulation, gives a direct (but inefficient) way to restrict the step size h in terms of the fixed (and known) quantity $\max\{|\lambda_{\min}|, |\lambda_{\max}|\}$, along the same lines for (1.0.12) as above. Root-finding tools can then be used with Theorems 1.0.1–1.0.2 to approximate the eigenvalues in $[\lambda_{\min}, \lambda_{\max}]$.

Problem (b) in page 1 on the other hand is quite different in general, since the indices $j_{\min}$ and $j_{\max}$ given in the problem formulation do not provide information about the size of $\lambda_{j_{\min}}$ or $\lambda_{j_{\max}}$. If one uses root-finding techniques with Theorem 1.0.2 to approximate the eigenvalues with the required indices, then one needs to have error control independent of

λ as one does not necessarily know its size, and therefore an ordinary differential equation solver cannot restrict the step size h to comply with the necessary accuracy, since λ is unknown. In such a situation, an approximation to (1.0.12) is unfortunately not enough, but rather a uniform approximation to (1.0.9) independent of λ is necessary.

Equally important for both problems (a) and (b), the development in this thesis of uniform approximations to (1.0.9), provides another contribution in that the approximations developed are independent of λ, which has the benefit that one does not have to reduce the step size h as a function of λ! This is in clear contrast with traditional solvers that only approximate (1.0.12), and decrease h for larger λ.

As mentioned briefly at the outset, the work in this dissertation spurs from a new concept named Fer streamers. This concept is shown to give rise to an approximation of (1.0.9) which holds uniformly for all (λ, t). This, as discussed above, permits the approximation of the eigenvalues and eigenfunctions of the Sturm–Liouville problems (1.0.1)–(1.0.2), in both variants (a) and (b). Indeed, as discussed carefully in Sections 1.1–1.2 below, the main advantage that separates this work via Fer streamers from every existing technique is that it is accompanied by error bounds which:

(i) hold uniformly for every eigenvalue, and,

(ii) can attain arbitrary high-order.

This work is composed of two parts, aggregated according to the regularity of the potential function, i.e., depending whether the potential is continuous and piecewise analytic:

$$q \in C^0\left([a,b],[q_{\min},q_{\max}]\right) \text{ is piecewise analytic}, \quad y_\lambda \in C^2\left([a,b],\mathbb{R}\right), \tag{1.0.13}$$

where many contributions in different directions exist throughout the literature, or whether the potential is absolutely integrable:

$$q \in \mathrm{L}^1\left([a,b],\mathbb{R}\right), \quad y_\lambda, y_\lambda' \in \mathrm{AC}\left([a,b],\mathbb{R}\right), \tag{1.0.14}$$

where results remain much more sparse.

- Firstly, in the main part of this thesis, this work considers the truncation, discretization, practical implementation and MATLAB software, of the new approach via Fer streamers for the classical setting with continuous and piecewise analytic potentials (1.0.13) in (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*).

For the classical setting (1.0.13), on the main part of this thesis, Section 1.1 compares the novel approach via Fer streamers (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*) with the seminal work of (Pruess, 1973; Paine and de Hoog, 1980; Marletta and Pryce, 1992; Pruess and Fulton, 1993) and the state-of-the-art work in (Ixaru, De Meyer and Berghe,

1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005; Iserles, 2004*a*; Ledoux and Daele, 2010; Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010; Moan, 1998), while Subsection 1.3.1 distills the various contributions that span (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*).

- Secondly, given that the techniques developed in (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*) are rather general, the second part of this thesis touches upon an extension of the new ideas that enabled the truncation of the new approach via Fer streamers, but instead for the general setting with absolutely integrable potentials (1.0.14) in (Ramos, 2014).

For the general setting (1.0.14), Section 1.2 and Subsection 1.3.2 discuss and summarize the contributions in (Ramos, 2014).

1.1 Relation to previous work

For continuous and piecewise analytic potentials (1.0.13), the present section serves to situate the novel contributions of (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*). In particular, they are related to and compared with the classical work of (Pruess, 1973; Paine and de Hoog, 1980; Marletta and Pryce, 1992; Pruess and Fulton, 1993), and the state-of-the-art work of (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005; Iserles, 2004*a*; Ledoux and Daele, 2010; Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010; Moan, 1998). Special emphasis is given to: *i*) geometric properties, *ii*) error estimates and number of evaluations of the potential, and, *iii*) volume of linear algebra, throughout the various methods.

1.1.1 Uniform but low-order error bounds

1.1.1.1 Piecewise Constant Methods

The Piecewise Constant Method (PCM) (Pruess, 1973; Paine and de Hoog, 1980; Marletta and Pryce, 1992; Pruess and Fulton, 1993) is among the earliest techniques used to approximate the eigenvalues of regular Sturm–Liouville problems. True for this day, it remains one of the few techniques mathematically guaranteed to approximate every eigenvalue uniformly well.

The underlying principle of the PCM consists in two approximations. First, approximate the eigenvalues of the original problem with those of a 'new' problem, defined as the original problem except q is replaced by a piecewise constant interpolation $\tilde{q} : [a, b] \to \mathbb{R}$. Second, compute the eigenvalues of the 'new' problem. The motivation is two-fold: on one hand, the first approximation is easily controlled with perturbation techniques, and, on the other hand, unlike the original problem, the 'new' problem is numerically tractable, up to prescribed tolerance.

That the PCM is sure to approximate all eigenvalues equally well, has long since been established (Pruess, 1973; Paine and de Hoog, 1980). More concretely, for each numerical mesh $c_0 := a < c_1 < \cdots < c_{m-1} < c_m := b$, $h_k := c_{k+1} - c_k$, $h_{\max} := \max\{h_0, \ldots, h_{m-1}\}$ such that $\tilde{q}|_{(c_k, c_{k+1})}$ is constant and interpolates q at some point in $[c_k, c_{k+1}]$, there exist error bounds in the uniform regime

$$h_{\max} \to 0^+, \text{ uniformly w.r.t. } \lambda \in \mathbb{R}, \tag{1.1.1}$$

where the constants in the big $\mathcal{O}$ notation are bounded independently of $\lambda \in \mathbb{R}$, that provide a convergence rate $d_1 h_{\max}^1$, where $d_1 > 0$ does not depend on $\lambda \in \mathbb{R}$. Among them, (Pruess, 1973, Theorem 1) controls the relative error with an error bound that yields uniform order 1 in the sense of (1.1.1). Another classical result is (Paine and de Hoog, 1980, Corollary 3.1), which controls the absolute error with an error bound that gives uniform order 1, again with respect to (1.1.1).

The uniform character of the error bounds for the PCM makes it an algorithm guaranteed to approximate every eigenvalue, at the same price. However, the convergence rate is rather low, which in practice means fine meshes and many function evaluations of q in order to construct $\tilde{q}$, where a popular choice is $\tilde{q}(t) := q((c_k + c_{k+1})/2)$ for t in (c_k, c_{k+1}). As function evaluations of q can be of considerable cost in practice, this creates a problem.

For this reason, there has been a lot of effort to develop algorithms with high convergence rate. Unfortunately, as discussed in the next subsection, the new results along this line of research have lost the uniform property of the error bounds in favor of a high convergence rate limited to 'small' or 'large' eigenvalues, so called 'asymptotic' rate.

1.1.2 High-order but non-uniform error bounds

1.1.2.1 Constant Perturbation Methods and modified Neumann integral series

The concept of 'asymptotic' order valid for 'small' or 'large' eigenvalues, alluded to in the previous subsection, was introduced first in (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005) to analyze the convergence of the Constant Perturbation Method (CPM). The guiding rule of the CPM consists of two truncations and one discretization. The first truncation is to approximate $q(c + ht)$, $t \in (0, 1)$, by the p-th degree polynomial of the Legendre series partial sum

$$\tilde{q}_p|_{(c,c+h)}(c + ht) := \sum_{j=0}^{p} q|_{(c,c+h)_j} P_j(2t - 1),$$

$$q|_{(c,c+h)_j} := (2j + 1) \int_0^1 q(c + ht) P_j(2t - 1) dt,$$

where $P_j(2t-1)$ denotes the j-th shifted Legendre polynomial. The second truncation is based on the PCM with $\lfloor \frac{2}{3}p \rfloor + 1$ corrections, rather than infinitely many. Together the two truncations form an approximation known as the CPM$[p, \lfloor \frac{2}{3}p \rfloor + 1]$. The ethos of the asymptotic order in (Ixaru, De Meyer and Berghe, 1997) and references that follow, is to investigate the truncation error of the CPM$[p, \lfloor \frac{2}{3}p \rfloor + 1]$ in the asymptotic regimes:

$$\lambda \text{ fixed and } h \to 0^+, \tag{1.1.2}$$

$$h \text{ fixed and } \lambda \to +\infty. \tag{1.1.3}$$

The approach corresponds directly to an analysis based on Taylor series or on asymptotic expansions. The first then provides some information about the behaviour of the approximations for 'small' eigenvalues whereas the second gives some insight into 'large' eigenvalues. Unfortunately, the underlying issue at play is that these asymptotic regimes are not well suited to study 'intermediary' eigenvalues, which require the control of (λ, h) instead of only either h (with fixed λ) or λ (with fixed h). Without being too precise, one of the difficulties here is that the power broker behind the scene is in fact

$$\begin{aligned} z_{\lambda,h} &:= \left(q|_{(c,c+h)_0} - \lambda \right) h^2, \\ \varpi_{\lambda,h} &:= 2\sqrt{-z_{\lambda,h}} = 2h\sqrt{\lambda - \frac{\int_c^{c+h} q\,(\xi)\,d\xi}{h}}, \end{aligned} \tag{1.1.4}$$

which appears, one way or another, as the argument of oscillatory functions (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005). For instance, as (Ixaru, De Meyer and Berghe, 1997, p. 294) puts it:

> "As λ is a free parameter (...) and we want to analyze the error behaviour at arbitrary λ, the whole range of $z_{\lambda,h}$'s has to be investigated. However, we can cover only two relevant extreme cases: $|z_{\lambda,h}|$ small and $z_{\lambda,h}$ large and negative."

These extreme cases correspond precisely to the asymptotic regimes (1.1.2)–(1.1.3) (Ixaru, De Meyer and Berghe, 1997, p. 295, p. 298). In particular, if one calls upon Taylor series in (1.1.2) as customary in the literature, then the factor λ and its powers populate the constants in the big $\mathcal{O}$ notation in the error bounds, making them useless for 'intermediary' or 'large' eigenvalues, unless one takes a prohibitively tiny step size, which is not an option in practice. Likewise, if one invokes asymptotic expansions in (1.1.3), one depends on $z_{\lambda,h} \ll -1$, making the error bounds unusable this time for 'small' or 'intermediary' eigenvalues. The truncation error of the CPM$[p, \lfloor \frac{2}{3}p \rfloor + 1]$ in these extreme cases is then controlled by (Ixaru, De Meyer and Berghe, 1997, p. 294):

- $r_{\lambda,p} h^{2p+2}$ w.r.t. (1.1.2), where $\lim_{\lambda \to +\infty} r_{\lambda,p} = +\infty$ and $\lim_{\lambda \to +\infty} r_{\lambda,p+1}/r_{\lambda,p} = +\infty$,
- $s_p h^p / \sqrt{\lambda}$ w.r.t. (1.1.3).

Finally, as $q|_{(c,c+h)_j}$, $j \in \{0, 1, \ldots, p\}$, are in general unavailable, they are approximated by quadrature, which forms the discretization step. For this, $q(c + h\cdot)$ is evaluated at p Gauss points in $(0, 1)$ (Ledoux, Daele and Berghe, 2004, p. 158) to form a quadrature with error h^{2p} for $j = 0$ and h^p for $j > 0$. Since $q|_{(c,c+h)_j}$ are always multiplied by h^2, this yields local error proportional to h^{p+2}, which caps the truncation bounds and yields the discretization bounds CPM$\{p + 2, p\}$ that behave as:

- $\tilde{r}_{\lambda,p} h^{p+2}$ w.r.t. (1.1.2), where $\lim_{\lambda\to+\infty} \tilde{r}_{\lambda,p} = +\infty$ and $\lim_{\lambda\to+\infty} \tilde{r}_{\lambda,p+1}/\tilde{r}_{\lambda,p} = +\infty$,
- $\tilde{s}_p h^p/\sqrt{\lambda}$ w.r.t. (1.1.3).

Since the introduction of the CPM (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005), there exist now different techniques that also attain asymptotic high-order. These include the modified Neumann integral series (Iserles, 2004*a*; Ledoux and Daele, 2010), which, as (Degani, 2004; Degani and Schiff, 2006) point out, are closely related to the CPM.

1.1.2.2 Modified or right correction Magnus integral series

Yet another approach is that of the modified or right correction Magnus Lie-group/Lie-algebra integral series (Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010). In passing, we note now that the last two methods are very similar in that (Ledoux, Daele and Berghe, 2010) extend the work by (Degani and Schiff, 2006) from $\lambda \gg q_{\max}$ to $\lambda \leq q_{\max}$ and propose different quadrature points for the (same) integrals.

This body of work presents several advancements when compared with (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005; Iserles, 2004*a*; Ledoux and Daele, 2010), such as the preservation of a certain geometric property (see Subsection 1.1.4). However, the analysis of the error bounds in (Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010) is limited also by the restrictions in the quote above from (Ixaru, De Meyer and Berghe, 1997, p. 294), which manifest along four fronts:

Firstly, and most importantly, these again are centered around the asymptotic regimes (1.1.2)–(1.1.3) with error bounds limited to 'small' or 'large' eigenvalues. Indeed, neither (1.1.2) nor (1.1.3) covers 'intermediary' eigenvalues. As a case in point, this is true in the truncation of the integral series as well as in the discretization of the multivariate integrals in (Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010). Secondly, the asymptotic regime (1.1.2) leads to quadrature estimates with 'large' constants in the big $\mathcal{O}$ notation, which are applicable only to 'small' eigenvalues. Thirdly, the asymptotic regime (1.1.3) leads to highly oscillatory multivariate quadrature, which is applicable only to 'large' eigenvalues and non-resonant integrals. Fourthly, both (1.1.2) and (1.1.3) lead to a localized decrease in function evaluations, for 'small' and 'large' eigenvalues.

To illustrate the first point above, note that the increase in global order from 6 to 8, with respect to the asymptotic regime (1.1.2), in the truncation of the integral series, in both (Degani and Schiff, 2006, Theorem 2) (c.f., (Degani, 2004, Theorem 7)) and (Ledoux, Daele and Berghe, 2010, Theorem 4.2), is built upon Taylor expansions (of oscillatory integrals) with coefficients that grow with λ. Thus, the $\mathcal{O}\big(h^{10}\big)$ terms in both (Degani, 2004, p. 35) and (Ledoux, Daele and Berghe, 2010, p. 759), grow with λ. As a consequence, the constants in the big $\mathcal{O}$ notation in every asymptotic estimate also grow with λ and are therefore applicable only to 'small' eigenvalues.

This is not inconsistent with the theorems in (Degani, 2004; Ledoux, Daele and Berghe, 2010), because they apply to the asymptotic regime (1.1.2), which, by definition, does not control the size of the constants in the big $\mathcal{O}$ notation. Indeed, the underlying issue at play is that the asymptotic regimes (1.1.2)–(1.1.3) are not well suited to control the 'intermediary' eigenvalues, which require the control of (λ, h) instead of only either h (with fixed λ) or λ (with fixed h).

As an example of the fourth point above, the localized decrease in function evaluations in the asymptotic regimes (1.1.2)–(1.1.3) is used to decrease the evaluations of the potential in (Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010) and can be traced to the quadrature estimates in the theoretical analysis in (Iserles, 2004*b*) for the univariate integral

$$h \int_0^1 \boldsymbol{A}_\lambda(ht) e^{it\varpi_{\lambda,h}} dt, \tag{1.1.5}$$

where $\varpi_{\lambda,h}$ is as in (1.1.4), which appears (often rewritten as a linear combination of different univariate integrals with different oscillatory kernels) in methods with global order greater than or equal to four. For example, the integral in (1.1.5) would appear after matching Eq. (3.1.1) together with Theorem 3.2.3, and a change of variables.

As (Iserles, 2004*b*) explains carefully, the quadrature estimates for (1.1.5) are different in different regimes: the non-oscillatory regime $|\varpi_{\lambda,h}| \ll 1$, the highly-oscillatory regime $\varpi_{\lambda,h} \gg 1$ and the intermediate regime $|\varpi_{\lambda,h}| \approx 1$. To be precise, with p quadrature points, the quadrature estimates are *i*) $\mathcal{O}\big(h^{2p+1}\big)$ for Gauss–Legendre and $\mathcal{O}\big(h^{2p-1}\big)$ for Gauss–Lobatto in the asymptotic regime with fixed λ and $h \to 0^+$, *ii*) $\mathcal{O}\big(h^{p+1}/\varpi_{\lambda,h}\big)$ for Gauss–Legendre and $\mathcal{O}\big(h^{p+1}/\varpi_{\lambda,h}^2\big)$ for Gauss–Lobatto in the asymptotic regime with fixed h and $\lambda \to +\infty$, and, *iii*) $\mathcal{O}\big(h^{p+1}\big)$ for both Gauss–Legendre and Gauss–Lobatto in the intermediary regime, where all bets are off and one needs to be extremely careful. This quantifies the localized decrease in function evaluations in the asymptotic regimes (1.1.2)–(1.1.3), for 'small' and 'large' eigenvalues. In addition,

$$\lim_{p\to+\infty} \frac{p+1}{2p+1} = \lim_{p\to+\infty} \frac{p+1}{2p-1} = \frac{1}{2}$$

also suggests a 50% localized increase in function evaluations for the univariate integral

(1.1.5), for the 'intermediary' eigenvalues.

The work in (Iserles, 2004*b*) also serves to illustrate the second point above because it quantifies the size of the constants in the big $\mathcal{O}$ notation in the quadrature estimates for (1.1.5). In fact, it is precisely to prevent the occurrence of 'large' constants in the big $\mathcal{O}$ notation that in that paper the quadrature estimates in the non-oscillatory regime $|\varpi_{\lambda,h}| \ll 1$ are different than the ones in the intermediary regime $|\varpi_{\lambda,h}| \approx 1$, and one of the reasons one needs to be extremely vigilant. This shows clearly that the quadrature estimates in the asymptotic regime (1.1.2) are valid only for 'small' eigenvalues.

To illustrate the third point above, it is important to recall that the quadrature estimates for (1.1.5) in the highly-oscillatory regime in (Iserles, 2004*b*) are built upon asymptotic expansions with $\varpi_{\lambda,h} \gg 1$ and to note that for fixed $\varpi_{\lambda,h}$, smaller h leads to larger λ. It is for these reasons that the quadrature estimates in the asymptotic regime (1.1.3) are valid only for 'large' eigenvalues and although high oscillation is an extremely effective device to decrease the quadrature error, it needs to be used with great care. Another reason for great caution with high oscillation, unique to the multivariate setting, is that the quadrature estimates in the asymptotic regime $\varpi_{\lambda,h} \gg 1$ are valid only in the absence of critical points and subject to a non-resonance condition (Iserles and Nørsett, 2006). In the context of the Lie-group/Lie-algebra integral series in (Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010; Ramos and Iserles, 2015; Ramos, 2015*a*), this is particularly important because the non-resonance condition is not satisfied in the bivariate integral

$$h^2 \int_0^1 \int_0^{t_1} \left[\boldsymbol{A}_\lambda(ht_2), \overline{\boldsymbol{A}_\lambda(ht_1)} \right] e^{i(t_2-t_1)\varpi_{\lambda,h}} dt_2 dt_1,$$

where $\varpi_{\lambda,h}$ is as in (1.1.4), which appears (sometimes rewritten as a linear combination of different bivariate integrals with different oscillatory kernels) in methods with global order greater than four. For instance, such bivariate integral would appear when combining Eq. (3.1.2) together with Theorem 3.2.3, and a change of variables.

Unlike the above results (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005; Iserles, 2004*a*; Ledoux and Daele, 2010; Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010), the next subsection presents a new set of ideas based on Fer streamers (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*) with which the whole range of $z_{\lambda,h}$'s is investigated. Unrestricted to the extreme cases described in the quote above from (Ixaru, De Meyer and Berghe, 1997, p. 294) that run through these approaches, the output is then the first algorithm that possesses error bounds which can attain arbitrary high-order and hold uniformly for all 'small', 'intermediary' and 'large' eigenvalues.

1.1.3 Uniform and high-order error bounds

The Fer streamers approach to Sturm–Liouville problems in the truncation of the integral series in the lead paper (Ramos and Iserles, 2015) and in the discretization of the multivariate integrals in (Ramos, 2015*a*), is virtually unique in the literature because it is based on the two uniform regimes

$$h_{\max} \to 0^+, \text{ uniformly w.r.t. } \lambda \in \left[q_{\max} - h_{\max}^{-2}, q_{\max} + h_{\max}^{-2}\right], \tag{1.1.6}$$

$$h_{\max} \to 0^+, \text{ uniformly w.r.t. } \lambda \in \left[q_{\max} + h_{\max}^{-2}, +\infty\right). \tag{1.1.7}$$

The only partial exception known to the author is the use of Magnus expansions by Moan in (Moan, 1998) who established a numerical method with global order four based on the single uniform regime

$$h \to 0^+, \text{ uniformly w.r.t. } \lambda \in \left[-h^{-2}, +h^{-2}\right]. \tag{1.1.8}$$

Indeed, as described below, the theory based on these two uniform regimes (1.1.6)–(1.1.7) is very different from the theory based on the two asymptotic regimes common throughout the literature (1.1.2)–(1.1.3), e.g., in (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005; Iserles, 2004*a*; Ledoux and Daele, 2010; Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010).

1.1.3.1 Moan's work with Magnus expansions

Moan's (1998) work, is based on four ideas: *i*) in formulating the Sturm–Liouville problem (1.0.1)–(1.0.2) in the Lie-group

$$\mathrm{SL}(2,\mathbb{R}) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a,b,c,d \in \mathbb{R} \text{ and } ad - bc = 1 \right\} \tag{1.1.9}$$

of two-by-two real matrices with determinant one, *ii*) in approximating the solution in $\mathrm{SL}(2,\mathbb{R})$ with the use of the Lie-algebra

$$\mathfrak{sl}(2,\mathbb{R}) := \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} : a,b,c \in \mathbb{R} \right\} \tag{1.1.10}$$

of two-by-two real matrices with zero trace and Magnus expansions, by calling upon (Iserles and Nørsett, 1999*b*), *iii*) in discretizing Magnus expansions, with discretization schemes put forth in (Iserles and Nørsett, 1999*b*), and, *iv*) in a clever summation of the discretized terms in order to avoid some of the issues that arise with large eigenvalues. In particular,

in his work, Moan (1998) established a numerical method with global order 4 which is able to approximate uniformly any eigenvalue within the bounded interval (1.1.8), where h denotes the step size.

Following Moan's (1998) work, Iserles, Munthe–Kaas, Nørsett and Zanna (2000) suggested in that paper a slightly different, but game-changing, approach: in short, switch the order of discretization and clever summation. This new idea, coined 'Magnus streamers', opened the door to truly fast computations of Magnus series in low-dimensional Lie algebras, making it an important contribution to the solution of matrix Lie-group linear differential equations. Unfortunately, when applied to the formulation of the Sturm–Liouville problem (1.0.1)–(1.0.2) in $\mathrm{SL}(2,\mathbb{R})$, the uniform approximation of Magnus streamers turn out to be prohibitively complex: there is nothing wrong with the summation, except that it is difficult to track the manner in which the magnitude of the eigenvalue influences the local and global error estimates.

In (Ramos and Iserles, 2015; Ramos, 2015*a*), we apply the aforementioned idea from (Iserles, Munthe–Kaas, Nørsett and Zanna, 2000) to Fer expansions' instead of Magnus expansions. The result is remarkable: by calling upon Fer expansions radius of convergence and recursive nature, it turns out that, unlike Magnus streamers, 'Fer streamers' lend themselves to uniform approximation of every eigenvalue and eigenfunction pair and exponentially growing order with increasing number of terms, making them a perfect tool in our endeavor!

1.1.3.2 Fer streamers

As mentioned above, unlike previous techniques (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005; Iserles, 2004*a*; Ledoux and Daele, 2010; Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010; Moan, 1998), Fer streamers (Ramos and Iserles, 2015; Ramos, 2015*a*) attain high-order without compromising the uniform property of the error bounds. Compared with the PCM (Pruess, 1973; Paine and de Hoog, 1980; Marletta and Pryce, 1992; Pruess and Fulton, 1993), they are thus also mathematically guaranteed to be uniformly precise but are not restricted to low-order.

At the heart of this advancement, Fer streamers are based on a completely new approach via Fer expansions, which is more algebraic in nature.

While a formal definition of Fer streamers will be given in Chapter 2, it is instructive to note here some of their salient features, namely, that one of the key points in Fer streamers is that they capitalize on the recursive nature of Fer expansions and exploit the low-dimensionality of a certain Lie-algebra to sum up the infinite sums in Fer expansions, in closed-form! This closed-form then makes it possible to bypass the approximation of each infinite sum in Fer expansions by its first partial sum — an approximation known to yield an error that grows with λ. As it turns out, by circumventing this approximation,

Fer streamers lead to an entirely new truncation and discretization of Fer expansions, with error bounds that hold equally well for all λ.

Another key element in Fer streamers is that their analysis calls upon Taylor series only for bounded functions with bounded derivatives with bounds independent of λ and abandons asymptotic expansions altogether for $z_{\lambda,h} \ll -1$.

Because of all this, the asymptotic regimes (1.1.2)–(1.1.3) do not even appear in the analysis of Fer streamers. Instead, once every algebraic feature is taken into account, it is the uniform regimes (1.1.6)–(1.1.7) that emerge naturally and it is with respect to these that Fer streamers derive error bounds, where the constants in the big $\mathcal{O}$ notation are bounded independently of $\lambda \in [q_{\max} - h_{\max}^{-2}, +\infty)$. In particular, given $p+1 \in \{4, 7, 10, 13, \ldots\}$, Fer streamers evaluate $q(a)$ and $q(c_k + h_k \cdot)$ at p points in $(0, 1]$ and yield total global error bounds $d_p h_{\max}^{p+1}$, with $d_p > 0$ independent of $\lambda \in [q_{\max} - h_{\max}^{-2}, +\infty)$.

1.1.4 Geometric integration

Apart from the different type of error bounds that separate these techniques, there is a geometric property left intact by some but not all, which pertains to the solution of the initial value problem (1.0.4)–(1.0.5), which appears in all of the above techniques, in essence because of the eigenvalue characterizations given by Theorems 1.0.1–1.0.2. In particular, there is a geometric feature intrinsic to (1.0.4)–(1.0.5) that should not go unnoticed: given that the matrix in (1.0.5) belongs to the Lie group (1.1.9) and the matrix in (1.0.4) lies in the Lie algebra (1.1.10), the solution possesses the geometric property:

$$\boldsymbol{Y}_\lambda([a, b]) \subseteq \mathrm{SL}(2, \mathbb{R}).$$

As documented in the literature, the preservation of this geometric feature leads to robust implementation of mismatch functions, such as $\lambda \mapsto \eta_\lambda$ in Theorem 1.0.1 and $\lambda \mapsto \theta_\lambda(b)$ in Theorem 1.0.2, which, as discussed above in pages 1–5, are invaluable tools to compute eigenvalues. It is important to note that regarding the above methods, the Fer streamers approach (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*) preserves this geometric feature as does the PCM (Pruess, 1973; Paine and de Hoog, 1980; Marletta and Pryce, 1992; Pruess and Fulton, 1993) and the modified or right correction Magnus (Iserles, 2004*b*; Degani, 2004; Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010), but not the CPM (Ixaru, De Meyer and Berghe, 1997, 1999; Ixaru, 2000; Ledoux, Daele and Berghe, 2004, 2005) nor the modified Neumann (Iserles, 2004*a*; Ledoux and Daele, 2010).

1.1.5 Computational complexity

The next subsubsections discuss how the computational complexity of the Fer streamers approach to Sturm–Liouville problems in (Ramos and Iserles, 2015; Ramos, 2015*a*) compares with alternative geometric integration techniques in the literature, with respect to:

i) number of steps in each numerical mesh, *ii*) error estimates and number of evaluations of the potential, and, *iii*) volume of linear algebra.

1.1.5.1 Number of steps in each numerical mesh

The truncation in (Ramos and Iserles, 2015) and the discretization in (Ramos, 2015*a*) are based on the following assumption:

Assumption 1.1.1. *The numerical mesh*

$$
\begin{aligned}
m &\in \mathbb{Z}^+, \\
c_0 := a < c_1 < \cdots < c_{m-1} < c_m &:= b, \\
h_k &:= c_{k+1} - c_k, \\
h_{\min} &:= \min_{k\in\{0,1,\ldots,m-1\}} \{h_k\}, \\
h_{\max} &:= \max_{k\in\{0,1,\ldots,m-1\}} \{h_k\},
\end{aligned}
$$

is such that

$$
\left.\begin{array}{r} k \in \{0,1,\ldots,m-1\} \\ t \in (c_k, c_{k+1}) \end{array}\right\} \quad \Longrightarrow \quad q(t) = \sum_{j=0}^{\infty} \frac{q^{(j)}\left(c_k^+\right)}{j!}(t-c_k)^j, \tag{1.1.11}
$$

$$
\lambda \geq q_{\min} \quad \Longrightarrow \quad h_{\max} \leq \frac{1}{\sqrt{q_{\max} - q_{\min}}}, \tag{1.1.12}
$$

$$
\lambda < q_{\min} \quad \Longrightarrow \quad h_{\max} \leq \frac{1}{\sqrt{q_{\max} - \lambda}}, \tag{1.1.13}
$$

$$
\frac{h_{\max}}{h_{\min}} \leq 2 \text{ } \textit{(this constant can be increased).} \tag{1.1.14}
$$

As discussed in (Ramos and Iserles, 2015), there exist Sturm–Liouville problems (1.0.1)–(1.0.2) where (1.1.13) does not need to be considered because there do not exist eigenvalues which are less than the minimum of the potential. For example, if the boundary conditions (1.0.2) are such that

$$
-y_\lambda'(b)y_\lambda(b) + y_\lambda'(a)y_\lambda(a) \geq 0 \tag{1.1.15}
$$

then

$$
\begin{aligned}
&-y_\lambda''(t) + q(t)y_\lambda(t) = \lambda y_\lambda(t) \\
&\Rightarrow \int_a^b \left(-y_\lambda''(t)y_\lambda(t) + q(t)\left(y_\lambda(t)\right)^2\right) dt = \lambda \int_a^b \left(y_\lambda(t)\right)^2 dt
\end{aligned}
$$

$$\Leftrightarrow \lambda = \frac{-y'_\lambda(b)y_\lambda(b) + y'_\lambda(a)y_\lambda(a) + \int_a^b \left((y'_\lambda(t))^2 + q(t)\,(y_\lambda(t))^2 \right) dt}{\int_a^b (y_\lambda(t))^2\, dt}$$

$$\Rightarrow \lambda \geq q_{\min}. \tag{1.1.16}$$

Important examples of boundary conditions (1.0.2) that satisfy (1.1.15) include zero Dirichlet

$$\alpha_1 \neq 0, \qquad \beta_1 \neq 0, \qquad \alpha_2 = \beta_2 = 0, \qquad y_\lambda(a) = y_\lambda(b) = 0$$

and zero Neumann

$$\alpha_1 = \beta_1 = 0, \qquad \alpha_2 \neq 0, \qquad \beta_2 \neq 0, \qquad y'_\lambda(a) = y'_\lambda(b) = 0$$

boundary conditions, but (1.1.16) is not true in general, as illustrated in (Ramos and Iserles, 2015). As an example, let

$$a = 0, \qquad b = \pi, \qquad (\forall t \in [0,\pi], q(t) = 0)\,, \qquad \alpha_1 = \alpha_2 \neq 0, \qquad \beta_1 = \beta_2 \neq 0$$

and consider the regular Sturm–Liouville problem in Liouville's normal form with self-adjoint separated boundary conditions

$$-y''_\lambda(t) = \lambda y_\lambda(t), \qquad t \in [0,\pi], \qquad y_\lambda(0) + y'_\lambda(0) = 0, \qquad y_\lambda(\pi) + y'_\lambda(\pi) = 0$$

with eigenvalues and eigenfunctions (normalized so that $\int_0^\pi (y_\lambda(t))^2\, dt = 1$) given in closed-form by

$$\lambda_j = \begin{cases} -1, & j = 0, \\ j^2, & j \in \mathbb{Z}^+, \end{cases}$$

$$y_{\lambda_j}(t) = \begin{cases} \dfrac{e^{-t}}{e^{-\frac{\pi}{2}}\sqrt{\sinh(\pi)}}, & j = 0, \\ \dfrac{j\cos(jt) - \sin(jt)}{\sqrt{\frac{\pi}{2}}\sqrt{j^2+1}}, & j \in \mathbb{Z}^+. \end{cases}$$

In this example, (1.1.16) does not hold true because the negative eigenvalue, $\lambda_0 = -1$, is strictly smaller than the minimum of the potential, $q_{\min} = 0$.

The need for assumption (1.1.13) is to prevent 'large' constants in the big $\mathcal{O}$ notation in the error estimates in both (Ramos and Iserles, 2015) and (Ramos, 2015*a*). In detail, if $\lambda < q_{\min}$ then the argument of certain hyperbolic cosines and sines is positive. If left unchecked, the argument becomes unbounded and the hyperbolic cosines and sines grow exponentially with the size of the argument, i.e., the constants in the big $\mathcal{O}$ notation become 'large'. Assumption (1.1.13) guarantees that in this case the positive argument is bounded by 2 and the hyperbolic cosines and sines are bounded by e^2, a 'small' constant.

Since the work in (Degani and Schiff, 2006, p. 423) assumes that $\lambda \gg q_{\max}$, this issue does not even arise. The work in (Ledoux, Daele and Berghe, 2010) does not assume $\lambda \geq q_{\min}$, and disregards this issue.

The use of assumption (1.1.12) is to enable an unhindered transition of the error estimates between the two uniform regimes (1.1.6)–(1.1.7). In addition, it quantifies the impact of the magnitude of the potential to the Fer streamers approach to Sturm–Liouville problems. The fact that the scale of the potential influences the step size is noted, but not quantified, in (Degani and Schiff, 2006, p. 416) and (Ledoux, Daele and Berghe, 2010, p. 761).

It is important to note that assumptions (1.1.12)–(1.1.13) require the knowledge of the minimum and the maximum of the potential. To be precise, only the knowledge of a lower bound to the minimum of the potential and an upper bound to its maximum are required, but the quality of the lower and upper bounds controls the maximum step size in view of assumptions (1.1.12)–(1.1.13). In (Degani and Schiff, 2006) the knowledge of the maximum of the potential is required in the sense that it focuses on the setting $\lambda \gg q_{\max}$. The work in (Ledoux, Daele and Berghe, 2010) does not use this information, since it does not control the positive argument of certain hyperbolic cosines and sines for $\lambda < q_{\min}$.

Assumption (1.1.14) controls the non-uniformity of the numerical mesh. As indicated in (1.1.14), the constant 2 can be increased, but it is important to understand that the non-uniformity of the numerical mesh is intrinsically related to the size of the constants in the big $\mathcal{O}$ notation in the error estimates.

1.1.5.2 Error estimates and number of evaluations of the potential

The discretization of the Fer streamers in (Ramos, 2015*a*), is made explicit with global orders 4, 7, 10 and 13, uniformly over the entire eigenvalue range (in the sense of the two uniform regimes (1.1.6) and (1.1.7)). Specifically, as proved in Theorem 3.4.3 below, the total global error in the Fer streamers approach to Sturm–Liouville problems (c.f., Definition 3.4.3) is controlled by the truncation global error (as defined in Definition 2.1.6 and upper bounded in Theorem 2.1.5) as well as by the discretization global error (as defined in Definition 3.4.2 and upper bounded in Theorem 3.4.2). The discretization in (Ramos, 2015*a*) with global orders 4, 7, 10, 13 requires 3 (two interior and one at the right boundary), 6 (five interior and one at the right boundary), 9 (eight interior and one at the right boundary), 12 (eleven interior and one at the right boundary) evaluations of the potential per mesh interval, respectively. Since the potential is continuous, this means that Fer streamers with global orders 4, 7, 10, 13 require $3m+1$, $6m+1$, $9m+1$, $12m+1$ evaluations of the potential for a mesh with m intervals, respectively.

The discretization in (Degani and Schiff, 2006, p. 422–429) is made explicit with global order 4 and 8, both in the asymptotic regime (1.1.2). It is also shown in (Degani and Schiff, 2006, Eq. 48) that the global error in that work is bounded in the asymptotic regime

(1.1.3). The discretization in (Degani and Schiff, 2006) with global order $4, 8$ requires $2, 4$ (interior) potential evaluations per mesh interval, which corresponds to $2m, 4m$ potential evaluations, respectively, for a mesh with m intervals.

As indicated above, (Ledoux, Daele and Berghe, 2010) extends the work by (Degani and Schiff, 2006) from $\lambda \gg q_{\max}$ to $\lambda \leq q_{\max}$ and suggests different potential evaluations. For global order $4, 8$ the discretization in (Ledoux, Daele and Berghe, 2010) instead uses $3, 5$ (one at each boundary and the rest in the interior) potential evaluations, respectively, per mesh interval. Given the potential is continuous, this translates into $2m + 1, 4m + 1$ potential evaluations, respectively, for a mesh with m intervals.

There is no analogue of Fer streamers with global orders 10 and 13 in (Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010).

As discussed in Subsection 1.1.2, the 50% localized increase in function evaluations from the asymptotic regimes (1.1.2)–(1.1.3) to the uniform regimes (1.1.6)–(1.1.7), is necessary for each univariate integral in order to control all 'small', 'intermediary' and 'large' eigenvalues.

1.1.5.3 Volume of linear algebra

The discretization schemes in (Ramos, 2015*a*) boil down to the quadrature of multivariate integrals of the form

$$h \int_0^1 \boldsymbol{Z}_\lambda(ht)dt, \tag{1.1.17}$$

$$h^2 \int_0^1 \int_0^{t_1} [\boldsymbol{Z}_\lambda(ht_2), \boldsymbol{Z}_\lambda(ht_1)]\, dt_2 dt_1, \tag{1.1.18}$$

$$h^3 \int_0^1 \int_0^{t_1} \int_0^{t_1} [\boldsymbol{Z}_\lambda(ht_3), [\boldsymbol{Z}_\lambda(ht_2), \boldsymbol{Z}_\lambda(ht_1)]]\, dt_3 dt_2 dt_1, \tag{1.1.19}$$

$$h^4 \int_0^1 \int_0^{t_1} \int_0^{t_1} \int_0^{t_1} [\boldsymbol{Z}_\lambda(ht_4), [\boldsymbol{Z}_\lambda(ht_3), [\boldsymbol{Z}_\lambda(ht_2), \boldsymbol{Z}_\lambda(ht_1)]]]\, dt_4 dt_3 dt_2 dt_1, \tag{1.1.20}$$

$$h^4 \int_0^1 \int_0^{t_1} \int_0^{t_1} \int_0^{t_2} [[\boldsymbol{Z}_\lambda(ht_4), \boldsymbol{Z}_\lambda(ht_2)], [\boldsymbol{Z}_\lambda(ht_3), \boldsymbol{Z}_\lambda(ht_1)]]\, dt_4 dt_3 dt_2 dt_1, \tag{1.1.21}$$

where $t \mapsto \boldsymbol{Z}_\lambda(ht)$ possesses a plethora of behaviour that varies with (λ, h). For instance, see (3.1.1)–(3.1.5). In detail, global order 4, 7, 10 and 13 in the uniform regimes (1.1.6)–(1.1.7) leads to the quadrature of multivariate integrals of the form (1.1.17), (1.1.17)–(1.1.18), (1.1.17)–(1.1.19) and (1.1.17)–(1.1.21), respectively.

The quadrature schemes in (Ramos, 2015*a*) are based on uniform approximations of $\boldsymbol{Z}_\lambda(ht)$ in $t \in [0, 1]$ by $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ with the property that (1.1.17)–(1.1.21), with $\boldsymbol{Z}_\lambda(h\cdot)$ replaced by $\tilde{\boldsymbol{Z}}_{\lambda,h}(\cdot)$, can be integrated exactly. The uniform approximations $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ are based on representations of $\boldsymbol{Z}_\lambda(ht)$ as finite sums such that each summand is a product of a bounded function with bounded derivatives and a trigonometric polynomial. An instance

with three summands and complex trigonometric polynomials can be found in Theorems 3.2.1 and 3.2.3, while an instance with four summands and real trigonometric polynomials can be found in Theorems 3.3.1 and 3.3.3. The uniform approximations $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ are build upon the polynomial interpolation of the aforementioned bounded functions with bounded derivatives.

Thus, the volume of linear algebra in the discretization schemes can be quantified (without accounting for reducing mechanisms such as free Lie algebras and Hall basis) by the number of terms in each integrand in (1.1.17)–(1.1.21) with $\tilde{\boldsymbol{Z}}_{\lambda,h}(\cdot)$ instead of $\boldsymbol{Z}_{\lambda}(h\cdot)$, which grows exponentially with i) base equal to the product between the number of summands in each representation of $\boldsymbol{Z}_{\lambda}(ht)$ times the number of points in the polynomial interpolation of the bounded functions with bounded derivatives, and, ii) exponent equal to the number of commutators in each integrand plus one.

Fortunately, the exponential growth of the number of terms in each integrand is heavily attenuated in the quadrature schemes in (Ramos, 2015*a*), since these require less interpolation points for the higher dimensional integrals than for the lower dimensional integrals. This, in turn, represents a significant saving in linear algebra. In detail, in the sense of the two uniform regimes (1.1.6)–(1.1.7),

- global order 4 requires $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ with 3 interpolation points for the univariate integral,
- global order 7 requires $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ with 6 interpolation points for the univariate integral and at most 3 interpolation points for the bivariate integral,
- global order 10 requires $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ with 9 interpolation points for the univariate integral, at most 6 interpolation points for the bivariate integral and at most 3 interpolation points for the trivariate integral, and,
- global order 13 requires $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ with 12 interpolation points for the univariate integral, at most 9 interpolation points for the bivariate integral, at most 6 interpolation points for the trivariate integral and at most 3 interpolation points for the quadrivariate integrals.

The "at most" feature described in the last three bullet points is made precise by the end of Subsection 3.3.2 and follows from Theorems 3.3.5, 3.3.6, 3.3.7 and 3.3.8.

The discretization of the methods in (Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010) with global order 4 and 8 with respect to (1.1.2), require the quadrature, with different $\boldsymbol{Z}_{\lambda}(ht)$, of (1.1.17) and (1.1.17)–(1.1.18), respectively. In detail, in (Degani and Schiff, 2006) global order $4, 8$ requires $\tilde{\boldsymbol{Z}}_{\lambda,h}(t)$ with $2, 4$ interpolation points for every multivariate integral, respectively, whereas in (Ledoux, Daele and Berghe, 2010) global order $4, 8$ instead uses $3, 5$ interpolation points for every multivariate integral, respectively.

There is no analogue of Fer streamers with uniform global orders 10 and 13 in (Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010).

In particular, (Degani and Schiff, 2006; Ledoux, Daele and Berghe, 2010) do not enjoy the heavy attenuation of the exponential growth of the number of terms in each integrand described above for the quadrature schemes in (Ramos, 2015*a*).

1.2 Broader settings

For absolutely integrable potentials (1.0.14), the current section discusses the ethos at the heart of the new contributions in (Ramos, 2014).

For clarity, let us focus on a relatively simple, but non-trivial example from (1.0.14). For instance, consider a potential q, analytic in (a,b), which belongs to $\mathrm{L}^1([a,b],\mathbb{R})$, but not to $\mathrm{L}^\infty([a,b],\mathbb{R})$. In other words, it is amenable throughout the interior of the interval, but blows up at least at one of the endpoints.

Given the task of approximating the eigensystem of (1.0.1)–(1.0.2) with such an unbounded potential, one is not granted access to the error bounds derived for, say, continuous and piecewise analytic potentials (1.0.13). For this reason, the traditional approach is then to:

(i) approximate the original potential q with a 'truncated' version $\tilde{q}$ analytic in $[a,b]$.

There are, however, several issues with this common approach. Firstly, it is often unclear how to construct $\tilde{q}$ given q, in order to compute the eigensystem up to prescribed tolerance. Secondly, since q and, very likely, its derivatives $q^{(j)}$ are unbounded near each singularity, is it very likely as well that $\tilde{q}$ and its derivatives $\tilde{q}^{(j)}$ although finite, are enormous in size, which leads to large constants in the big $\mathcal{O}$ notation in the error bounds with $\tilde{q}$ and imposes small step sizes in the numerical mesh. To counter such issues, the practice is then to use an heuristic mesh selection algorithm that refines the mesh for $\tilde{q}$ severely near singularities of q, but results in larger step sizes away from them, and updates $\tilde{q}$ interactively. Contrary to this common approach, the work in (Ramos, 2014) starts anew and pursues a different avenue of research which is to:

(ii) work directly with the original potential q without truncation.

Following this new line of research, as a first step, (Ramos, 2014) reexamines and generalizes the work in (Ramos and Iserles, 2015). In particular, it is found in (Ramos, 2014) that the error bounds in (Ramos and Iserles, 2015) retain the same uniform and high-order stellar features either in the vanilla setting (1.0.13) or in the general case (1.0.14), which places Fer streamers in a unique position to pursue the rigorous approximation of the eigensystem of (1.0.1)–(1.0.2) in much broader settings as well.

1.3 Outline and contributions of the thesis

As examined above in Sections 1.1–1.2, the paramount property that distinguishes this novel approach based on Fer streamers from previously existing techniques rests in the fact that it possesses error bounds which: *i*) hold uniformly over the entire eigenvalue range, and, *ii*) can attain arbitrary high-order.

In line with the presentation above, the work in this thesis on the novel approach via Fer streamers to regular Sturm–Liouville problems (1.0.1)–(1.0.2), is organized depending on the regularity of the potential, i.e., on either (1.0.13) or (1.0.14).

Firstly, embodying the main novelties in this work, Chapters 2, 3, 4, 5 present, respectively, the work in (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*,*c*), the contributions in each being summarized below in Subsubsections 1.3.1.1, 1.3.1.2, 1.3.1.3, 1.3.1.4.

Secondly, motivated by the generality of the techniques used in the new approach, Chapter 6 discusses the work in (Ramos, 2014), which extends the scope of the approach in (Ramos and Iserles, 2015), from (1.0.13) to (1.0.14), the novelty of which is condensed in Subsubsection 1.3.2.1 below.

1.3.1 Continuous and piecewise analytic potentials

1.3.1.1 A new truncation: from Fer expansions to Fer streamers (Chapter 2)

In (Ramos and Iserles, 2015), we put forth a new set of error bounds to approximate the eigenvalues and eigenfunctions of regular Sturm–Liouville problems, in Liouville's normal form, defined on compact intervals (1.0.1), with continuous and piecewise analytic potentials (1.0.13) and self-adjoint separated boundary conditions (1.0.2).

The point of departure in (Ramos and Iserles, 2015) is to interpret the problem setting in a Lie-group/Lie-algebra formalism and to capitalize on the low-dimensionality of the Lie algebra to rewrite any analytic function of any commutator matrix in a very useful form. This basic idea was then melded with Fer expansions to produce a new concept called 'Fer streamers', setting the stage for a non-standard truncation of Fer expansions.

This new concept was nurtured throughout (Ramos and Iserles, 2015) and resulted in an approximation, which: *i*) does not impose any restriction on the step size for eigenvalues which are greater than or equal to a certain constant, *ii*) requires only a mild restriction on the step size for the remaining finite number of eigenvalues, *iii*) can attain any convergence rate, which grows exponentially with the number of terms, and is uniform for every eigenvalue, and *iv*) lends itself to a clear understanding of the manner in which the potential affects the local and global truncation errors.

1.3.1.2 Retaining Fer streamers' properties under discretization (Chapter 3)

The following paper (Ramos, 2015*a*) covers the discretization of the novel approach to the computation of regular Sturm–Liouville problems via Fer streamers, introduced in (Ramos and Iserles, 2015). The motivation to discretize the novel approach via Fer streamers stems from the local and global truncation bounds in (Ramos and Iserles, 2015) which guarantee large step sizes uniform over the entire eigenvalue range and tight error estimates uniform for every eigenvalue. The work in (Ramos, 2015*a*) shows how to retain these advantageous features under discretization, which is made explicit for global orders 4, 7, 10 and 13. The interplay between the truncation and the discretization in the approach by Fer streamers is also carefully quantified with total error bounds in (Ramos, 2015*a*).

1.3.1.3 Decreasing the volume of linear algebra in Fer streamers (Chapter 4)

While (Ramos and Iserles, 2015; Ramos, 2015*a*) focused on developing Fer streamers with uniform and high-order truncation and discretization error bounds, the paper (Ramos, 2015*b*) instead explains how to capitalize on a reduced Hall basis to yield an efficient implementation, by decreasing the volume of linear algebra in this new approach. Once again, special emphasis is given to Fer streamers with uniform global orders 4, 7, 10 and 13.

1.3.1.4 Fer streamers' MATLAB package (Chapter 5)

The work in (Ramos and Iserles, 2015; Ramos, 2015*a*,*b*) has now been realized in the form of a MATLAB package, with uniform global orders 4, 7, 10 and 13, presented for the first time in (Ramos, 2015*c*). Apart from serving the practitioner, this MATLAB package, illustrates also the power of the results via Fer streamers.

1.3.2 Absolutely integrable potentials

1.3.2.1 A generalized truncation (Chapter 6)

In (Ramos, 2014), we reexamine and generalize the results of (Ramos and Iserles, 2015). In particular, (Ramos, 2014) proves that the Fer streamers' approximation developed in (Ramos and Iserles, 2015) retains its uniform and high-order useful properties either in the original setting (1.0.13) or in the general case (1.0.14).

Although the basic idea is the concept of Fer streamers introduced in (Ramos and Iserles, 2015), this general case presents several subtleties which need to be identified and addressed. In particular, we identify four nested classes of potentials which require different treatment, e.g., different inequalities, different restrictions on the step size, different selection criteria on the numerical mesh, different flows or different non-linear characterizations of the eigenvalues. For example, the last three points are especially important whenever

the potential is absolutely integrable but not in $\mathrm{L}^p([a,b],\mathbb{R})$, $p \in (1,\infty]$, since *i*) the mesh points, which are not boundary points, have to be Lebesgue points of the potential, and *ii*) if the left boundary point is not a Lebesgue point of the potential then the flow needs to be separated into 'positive' and 'negative' parts and the non-linear characterization of the eigenvalues needs to be changed.

Chapter 2

A new truncation: from Fer expansions to Fer streamers

As explained carefully in Chapter 1, the chief advantage of Fer streamers over previous approaches, lies on the fact that they are accompanied by error bounds which:

(i) hold uniformly over the entire eigenvalue range, in the sense of (1.1.6)–(1.1.7), and,

(ii) can attain arbitrary high-order.

The present chapter presents the first contribution towards the development of the novel approach via Fer streamers introduced in (Ramos and Iserles, 2015), for regular Sturm–Liouville problems, in Liouville's normal form, defined on compact intervals (1.0.1), with self-adjoint separated boundary conditions (1.0.2), and continuous and piecewise analytic potentials (1.0.13).

The new approach based on Fer streamers' (Ramos and Iserles, 2015) consists in a three-step procedure, which is centered on Assumption 1.1.1. As touched upon in Chapter 1, the first step is based in formulating the Sturm–Liouville problem (1.0.1)–(1.0.2) in the Lie-group (1.1.9) of two-by-two real matrices with determinant one, and in approximating uniformly, with respect to (1.1.6)–(1.1.7), the solution in the Lie-group with the use of the Lie-algebra (1.1.10) of two-by-two real matrices with zero trace. In particular, towards this end, according to the two cases distinguished in (1.1.12) and (1.1.13), the eigenvalue interval $[q_{\max} - h_{\max}^{-2}, +\infty)$ is divided into two pieces

$$\lambda \in \left[q_{\max} - h_{\max}^{-2}, q_{\max} + h_{\max}^{-2}\right] \cup \left[q_{\max} + h_{\max}^{-2}, +\infty\right) \tag{2.0.1}$$

and we approximate the solution of the initial value problem (1.0.4)–(1.0.5) in the two uniform regimes (1.1.6)–(1.1.7).

Our main ideas lie precisely in the development of these uniform expansions. In particular, we proceed by recalling Fer expansions below in Subsection 2.1.1 and observing that

the error in the standard truncation of Fer expansions deteriorates with increasing values of λ. This, at first glance, suggests that Fer expansions are not a useful tool to increase the step size in the presence of large eigenvalues, but nothing could be further from the truth! Indeed, as shown below in Subsection 2.1.2, it is possible to truncate Fer expansions in an alternative manner, with what we call Fer streamers, which, to all intents and purposes, do not impede the step size and yield error estimates with exponentially growing order with increasing number of terms, which also single out the role of the potential!

As explained earlier in Chapter 1, the second and third steps are standard: approximate the eigenvalues via root-finding together with either $\lambda \mapsto \eta_\lambda$ in Theorem 1.0.1 or $\lambda \mapsto \theta_\lambda(b)$ in Theorem 1.0.2. This, as explained before in pages 1–5, can be achieved by approximating $\boldsymbol{Y}_\lambda(c_{k+1})$ uniformly with Fer streamers, and solving the resulting approximate equations of $\lambda \mapsto \eta_\lambda$ and $\lambda \mapsto \theta_\lambda(b)$ with the use of a root-finding algorithm. Having approximated the eigenvalues, one can then approximate the eigenfunctions with (1.0.10)–(1.0.11), which again are based on these uniform approximations.

2.1 Fer expansions and streamers

We embark in this section upon the core of our argument and the essence of the novelty of its contribution, namely the elaboration of an approximation of (1.0.4)–(1.0.5) in the two uniform regimes (1.1.6) and (1.1.7). We note that it is the uniform character of our approximations which makes them a very useful tool in our endeavor to approximate small, medium or large eigenvalues of Sturm–Liouville problems.

In the following Subsection 2.1.1, we recall Fer expansions and observe that they provide an amenable closed-form representation of the exact solution of (1.0.4)–(1.0.5), with two important properties: Firstly, Fer expansions are valid whenever the potential is piecewise analytic, a feature independent of any eigenvalue. Secondly, Fer expansions are naturally defined via a recurrence relation.

It is then, in the subsequent Subsection 2.1.2, that we establish, under the mild conditions, (1.1.11), (1.1.12), (1.1.13) and (1.1.14), that those two properties pave the way to the uniform approximation of what we call Fer streamers: exact closed-form expressions which we devise for each of the terms appearing in Fer expansions. This, in turn, is shown to yield a uniform approximation of $\boldsymbol{Y}_\lambda(c_{k+1})$.

2.1.1 Fer expansions

For 'small' eigenvalues, it is possible to solve (1.0.4)–(1.0.5) by calling upon the following definitions and theorem from (Fer, 1958; Iserles, 1984, Theorem 3; Iserles, Munthe–Kaas, Nørsett and Zanna, 2000, p. 267–270).

Definition 2.1.1. *Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathfrak{sl}(2,\mathbb{R})$, and define the exponential, the adjoint representation, and the derivative of the adjoint representation (also referred to as the Lie bracket) as*

$$\begin{aligned}
\rho\big(\boldsymbol{X}\big) &:= 2\sqrt{-\det\left(\boldsymbol{X}\right)},\\
\exp\left(\boldsymbol{X}\right) &:= \sum_{j=0}^{\infty}\frac{\boldsymbol{X}^j}{j!} = \cosh\left(\frac{\rho(\boldsymbol{X})}{2}\right)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{\sinh\left(\frac{\rho(\boldsymbol{X})}{2}\right)}{\frac{\rho(\boldsymbol{X})}{2}}\boldsymbol{X},\\
\mathrm{Ad}_{\exp(\boldsymbol{X})}\boldsymbol{Y} &:= \exp\left(\boldsymbol{X}\right)\boldsymbol{Y}\exp\left(-\boldsymbol{X}\right),\\
\mathrm{ad}_{\boldsymbol{X}}\boldsymbol{Y} &=: [\boldsymbol{X},\boldsymbol{Y}] := \boldsymbol{X}\boldsymbol{Y} - \boldsymbol{Y}\boldsymbol{X}.
\end{aligned}$$

Remark 2.1.1 (Ramos and Iserles, 2015; Ramos, 2015*a*)**.** *Note that the exponential is in* $\mathrm{SL}(2,\mathbb{R})$ *and that the adjoint representation and the derivative of the adjoint representation are in* $\mathfrak{sl}(2,\mathbb{R})$*. It is also important to note that the exponential map from the Lie algebra* $\mathfrak{sl}(2,\mathbb{R})$ *to the Lie group* $\mathrm{SL}(2,\mathbb{R})$ *possesses the well-known closed-form in Definition 2.1.1 (Iserles, Munthe–Kaas, Nørsett and Zanna, 2000, Section 8).*

Definition 2.1.2. *Let $l \in \mathbb{Z}^+$ and $t \in [c_k, c_{k+1}]$, and define*

$$\boldsymbol{B}_{\lambda,0}(c_k,t) := \begin{bmatrix}0 & 1\\ q(t)-\lambda & 0\end{bmatrix}, \tag{2.1.1}$$

$$\boldsymbol{D}_{\lambda,0}(c_k,t) := \int_{c_k}^{t}\boldsymbol{B}_{\lambda,0}(c_k,\xi)d\xi = (t-c_k)\begin{bmatrix}0 & 1\\ \frac{\int_{c_k}^{t}q(\xi)d\xi}{t-c_k}-\lambda & 0\end{bmatrix}, \tag{2.1.2}$$

$$\boldsymbol{B}_{\lambda,l}(c_k,t) := \sum_{j=1}^{\infty}(-1)^j\frac{j}{(j+1)!}\mathrm{ad}^j_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{B}_{\lambda,l-1}(c_k,t), \tag{2.1.3}$$

$$\boldsymbol{D}_{\lambda,l}(c_k,t) := \int_{c_k}^{t}\boldsymbol{B}_{\lambda,l}(c_k,\xi)d\xi. \tag{2.1.4}$$

Remark 2.1.2 (Ramos and Iserles, 2015)**.** *Observe that $\boldsymbol{B}_{\lambda,0}(c_k,t)$, $\boldsymbol{D}_{\lambda,0}(c_k,t)$, $\boldsymbol{B}_{\lambda,1}(c_k,t)$, $\boldsymbol{D}_{\lambda,1}(c_k,t),\ldots \in \mathfrak{sl}(2,\mathbb{R})$. This was recognized in Zanna's (1996) work (see the historical reference by Iserles, Munthe–Kaas, Nørsett and Zanna, p. 267–270), and will go a long way to retain the geometric feature described in Subsection 1.1.4.*

Theorem 2.1.1 (Fer, 1958; Iserles, 1984, Theorem 3; Iserles, Munthe–Kaas, Nørsett and Zanna, 2000, p. 267–270)**.** *If* (1.1.11) *holds true and* $l \in \mathbb{Z}^+$*, then*

$$\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1}) = h_{\max} \begin{bmatrix} \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) \end{bmatrix}, \qquad \textit{with respect to } (1.1.2), \tag{2.1.5}$$

$$\boldsymbol{D}_{\lambda,l}(c_k, c_{k+1}) = h_{\max}^{4\times 2^{l-1}-1} \begin{bmatrix} \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) \end{bmatrix}, \qquad \textit{with respect to } (1.1.2), \tag{2.1.6}$$

where some constants in the big $\mathcal{O}$ *notation grow with increasing* λ*, and the solution of* (1.0.4) *is given by the Fer expansions*

$$\boldsymbol{Y}_\lambda(c_{k+1}) = \left(e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\boldsymbol{D}_{\lambda,1}(c_k,c_{k+1})} e^{\boldsymbol{D}_{\lambda,2}(c_k,c_{k+1})} \cdots \right) \boldsymbol{Y}_\lambda(c_k). \tag{2.1.7}$$

Although Theorem 2.1.1 provides a closed-form representation of the exact solution of (1.0.4), it is not clear in practice how to evaluate or approximate (2.1.7). In particular, Theorem 2.1.1 does not provide a practical means to evaluate or approximate the infinite series (2.1.3). This state of affairs was partially resolved in (Zanna, 1998). The methodology in (Zanna, 1998) consists in two levels of truncation: one in the infinite product of exponentials in (2.1.7), and one in each infinite sum (2.1.3). Specifically, Zanna's (1998) work succeeds in approximating the exact solution by calling upon (2.1.5)–(2.1.6) to discard all except the very first exponentials in the infinite product (2.1.7), and by a careful estimation of each summand to discard all except the very first terms in each infinite sum (2.1.3). This procedure works exceedingly well, but only for 'small' values of $|\lambda|$, since, as indicated above, some of the constants in the big $\mathcal{O}$ notation in Theorem 2.1.1 increase with growing λ.

2.1.2 Fer streamers

As pointed out, for 'medium' or 'large' eigenvalues, however, the two-stage truncation procedure described in the previous subsection breaks down, and leads to catastrophic results. Indeed, it is possible to see that in the aforementioned procedure: *i*) it is only feasible to solve for eigenvalues in a compact interval $h_{\max}^2|\lambda| \leq 1$, as opposed to an unbounded interval, and, *ii*) the error bounds deteriorate quite considerably, or completely, whenever $h_{\max}^2|\lambda| \approx 1$.

We now address this issue by proposing a non-standard truncation of Fer expansions, which consists in one less level of truncation. Our point of departure is what we call Fer streamers: exact closed-form expressions which we devise for each infinite sum (2.1.3).

With these closed-form expressions at hand, we are left only with the truncation of the infinite product (2.1.7), and we proceed by investigating the size of Fer streamers, in the two uniform regimes (1.1.6) and (1.1.7).

The result, is a uniform approximation of $\boldsymbol{Y}_\lambda(c_{k+1})$, which, under the mild conditions (1.1.11), (1.1.12), (1.1.13) and (1.1.14), provides a means to estimate any eigenvalue with little or no restriction on the step size! Moreover, our proposed uniform approximation retains the same, albeit slower, type of exponential growth in order!

2.1.2.1 Closed-form expressions

The current subsubsection concerns the exact sum in closed-form of the infinite series in (2.1.3), which is achieved in Theorem 2.1.3 and illustrated for a particular example in Remark 2.1.5.

Definition 2.1.3. *For every* $\boldsymbol{X} \in \mathfrak{sl}(2,\mathbb{R})$ *and* $\boldsymbol{x} \in \mathbb{R}^{3\times 1}$*, let*

$$\boldsymbol{\pi}(\boldsymbol{X}) := \begin{bmatrix} [\boldsymbol{X}]_{1,1} \\ [\boldsymbol{X}]_{1,2} \\ [\boldsymbol{X}]_{2,1} \end{bmatrix}, \qquad \boldsymbol{\pi}^{-1}(\boldsymbol{\pi}(\boldsymbol{X})) = \boldsymbol{X},$$

$$\boldsymbol{\pi}^{-1}(\boldsymbol{x}) := \begin{bmatrix} [\boldsymbol{x}]_{1,1} & [\boldsymbol{x}]_{2,1} \\ [\boldsymbol{x}]_{3,1} & -[\boldsymbol{x}]_{1,1} \end{bmatrix}, \qquad \boldsymbol{\pi}(\boldsymbol{\pi}^{-1}(\boldsymbol{x})) = \boldsymbol{x},$$

$$\mathscr{C}_{\boldsymbol{X}} := \begin{bmatrix} 0 & -[\boldsymbol{X}]_{2,1} & [\boldsymbol{X}]_{1,2} \\ -2[\boldsymbol{X}]_{1,2} & 2[\boldsymbol{X}]_{1,1} & 0 \\ 2[\boldsymbol{X}]_{2,1} & 0 & -2[\boldsymbol{X}]_{1,1} \end{bmatrix}.$$

Remark 2.1.3 (Ramos, 2015*a*)**.** *It should not go unnoticed in Definition 2.1.3 that* $\boldsymbol{\pi}$ *embeds* $\mathfrak{sl}(2,\mathbb{R})$ *in* $\mathbb{R}^{3\times 1}$*. This embedding is possible because* $\dim \mathfrak{sl}(2,\mathbb{R}) = 3$ *and it is at the core of the novel contributions in (Ramos and Iserles, 2015) reported in the current chapter. In the following chapter (Ramos, 2015a), it is key to develop the discretization schemes in Section 3.3.*

Theorem 2.1.2 (Ramos and Iserles, 2015)**.** *If* $l \in \mathbb{Z}^+$ *and* $\boldsymbol{X}, \boldsymbol{Y} \in \mathfrak{sl}(2,\mathbb{R})$*, then*

$$\boldsymbol{\pi}(\mathrm{ad}_{\boldsymbol{X}}\boldsymbol{Y}) = \mathscr{C}_{\boldsymbol{X}}\boldsymbol{\pi}(\boldsymbol{Y}), \qquad \mathscr{C}_{\boldsymbol{X}}^{2l-1} = \rho^{2l-2}(\boldsymbol{X})\mathscr{C}_{\boldsymbol{X}},$$
$$\mathrm{ad}_{\boldsymbol{X}}\boldsymbol{Y} = \boldsymbol{\pi}^{-1}(\mathscr{C}_{\boldsymbol{X}}\boldsymbol{\pi}(\boldsymbol{Y})), \qquad \mathscr{C}_{\boldsymbol{X}}^{2l} = \rho^{2l-2}(\boldsymbol{X})\mathscr{C}_{\boldsymbol{X}}^{2}.$$

Proof. The assertions on the left follow by straightforward computation, and the ones on the right follow by induction from

$$\mathscr{C}_{\boldsymbol{X}}^3 = \rho^2(\boldsymbol{X})\mathscr{C}_{\boldsymbol{X}},$$

which itself follows from Cayley–Hamilton's theorem since

$$\det\left(t\boldsymbol{I}_3 - \mathscr{C}_{\boldsymbol{X}}\right) = t^3 - \rho^2(\boldsymbol{X})t.$$

□

Definition 2.1.4. *Let*

$$\psi(z) := \sum_{j=1}^{\infty}(-1)^j \frac{j}{(j+1)!}z^j = -\frac{e^{-z}(e^z - 1 - z)}{z}$$

and

$$\varphi(z) := \frac{\psi(z) - \psi(-z)}{2z} = -\sum_{j=0}^{\infty}\frac{2j+1}{(2j+2)!}z^{2j} = \frac{\cosh(z) - 1 - z\sinh(z)}{z^2},$$
$$\phi(z) := \frac{\psi(z) + \psi(-z)}{2z^2} = \sum_{j=0}^{\infty}\frac{2j+2}{(2j+3)!}z^{2j} = \frac{z\cosh(z) - \sinh(z)}{z^3}.$$

Remark 2.1.4 (Ramos and Iserles, 2015)**.** *In the sequel, it will be vital to observe that both φ and ϕ are bounded along the imaginary axis:*

$$\varphi(ix) = \sum_{j=0}^{\infty}(-1)^{j+1}\frac{2j+1}{(2j+2)!}x^{2j} = \left(\frac{1-\cos(x)}{x} - \sin(x)\right)\frac{1}{x},$$
$$\phi(ix) = \sum_{j=0}^{\infty}(-1)^{j}\frac{2j+2}{(2j+3)!}x^{2j} = \left(\frac{\sin(x)}{x} - \cos(x)\right)\frac{1}{x^2}.$$

We name the exact closed-form expressions which appear in the following Theorem, as Fer streamers.

Theorem 2.1.3 (Ramos and Iserles, 2015)**.** *If (1.1.11) holds true, $l \in \mathbb{Z}^+$ and $t \in [c_k, c_{k+1}]$, then the infinite series appearing in the terms of the Fer expansions of the initial value problem (1.0.4)–(1.0.5) in Definition 2.1.2 are given in closed-form by the 'Fer streamers'*

$$\begin{aligned}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right) = {}& \varphi\left(\rho\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\ &+ \phi\left(\rho\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right),\end{aligned}$$

which, equivalently, can be written as

$$\begin{aligned}\boldsymbol{B}_{\lambda,l}(c_k,t) = \varphi\left(\rho\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\operatorname{ad}_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{B}_{\lambda,l-1}(c_k,t)\\ + \phi\left(\rho\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\operatorname{ad}^2_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{B}_{\lambda,l-1}(c_k,t).\end{aligned}$$

Proof. See Section 2.3. □

Remark 2.1.5 (Ramos and Iserles, 2015)**.** *As an example, since*

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right) = \begin{bmatrix}0\\1\\q(t)-\lambda\end{bmatrix}, \quad \boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) = (t-c_k)\begin{bmatrix}0\\1\\\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda\end{bmatrix},$$

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) = 2(t-c_k)\sqrt{\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda}, \tag{2.1.8}$$

we have that Theorem 2.1.3 yields (see Section 2.4)

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right) = \begin{bmatrix}\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\frac{q(t)-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)^2\\ -2\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\frac{q(t)-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)^3\\ \frac{1}{2}\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\frac{q(t)-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)\end{bmatrix}.$$

Remark 2.1.6. *As touched upon in Section 1.1, (1.1.4), a quantity which intertwines λ and h, always appears in the numerical solution of Sturm–Liouville problems, one way or another, as the argument of oscillatory functions. Here, with Fer streamers, it relates to (2.1.8) as it takes the form*

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c,c+h\sigma)\right) = i\sigma\varpi_{\lambda,h}\sqrt{\frac{\lambda-\frac{\int_c^{c+h\sigma}q(\xi)d\xi}{h\sigma}}{\lambda-\frac{\int_c^{c+h}q(\xi)d\xi}{h}}}, \quad \sigma\in[0,1].$$

In view of Remark 2.1.4, Remark 2.1.5 places it once again as the argument of oscillatory functions, this time those being

$$x\mapsto\varphi(ix), \quad x\mapsto-2\phi(ix), \quad x\mapsto\frac{1}{2}\phi(ix)(ix)^2.$$

In particular, if both $\varpi_{\lambda,h}$ and λ are large and positive then $\boldsymbol{B}_{\lambda,1}(c,c+h\sigma)$ is an highly

oscillatory matrix function for $\sigma \in [0,1]$. Central to this work, this realization arises from having written the infinite series in Definition 2.1.2 in closed-form via Fer streamers in Theorem 2.1.3, since otherwise it would remain unnoticed.

2.1.2.2 Error estimates

With Fer streamers from Theorem 2.1.3 in hand, the present subsubsection puts forth a uniform approximation of $\boldsymbol{Y}_{\lambda}(c_{k+1})$, denoted by $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$, the approximation properties of which are unveiled below in Corollary 2.1.1.

Definition 2.1.5. *Let*

$$\delta_{|q'|} := \max_{k\in\{0,1,\ldots,m-1\}} \max_{t\in(c_k,c_{k+1})} \left\{|q'(t)|\right\} = \|q'\|_{\mathrm{L}^{\infty}([a,b],\mathbb{R})}.$$

Theorem 2.1.4 (Ramos and Iserles, 2015)**.** *If Assumption 1.1.1 holds true, $l \in \mathbb{Z}^{+}$ and $t \in [c_k, c_{k+1}]$, then, in the uniform regime (1.1.6), it follows that*

$$e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\cdots e^{\boldsymbol{D}_{\lambda,0}(a,c_1)} = \begin{bmatrix} \mathcal{O}\left(1\right) & \mathcal{O}\left(h_{\max}\right) \\ \mathcal{O}\left(h_{\max}^{-1}\right) & \mathcal{O}\left(1\right) \end{bmatrix},$$

$$\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right) = \delta_{|q'|}^{2^{l-1}} h_{\max}^{3\times 2^{l-1}-1} \begin{bmatrix} \mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^{2}\right) & \mathcal{O}\left(1\right) \end{bmatrix}^{\top},$$

and, in the uniform regime (1.1.7), it follows that

$$e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\cdots e^{\boldsymbol{D}_{\lambda,0}(a,c_1)} = \begin{bmatrix} \mathcal{O}\left(1\right) & \frac{\mathcal{O}(1)}{\sqrt{\lambda-q_{\max}}} \\ \mathcal{O}\left(1\right)\sqrt{\lambda-q_{\max}} & \mathcal{O}\left(1\right) \end{bmatrix},$$

$$\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right) = \delta_{|q'|}^{2^{l-1}} h_{\max}^{2^{l}} (\lambda-q_{\max})^{-\frac{2^{l-1}-1}{2}} \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda-q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda-q_{\max}} & \mathcal{O}\left(1\right) \end{bmatrix}^{\top},$$

where the constants in the big $\mathcal{O}$ notation can be bounded independently of λ.

Proof. See Section 2.4. □

It is insightful to compare between the first estimates for the size of $\boldsymbol{D}_{\lambda,l}(c_k,t)$, $l \geq 1$, derived without Fer streamers in Theorem 2.1.1 for the asymptotic regime (1.1.2), and the second estimates — for the same quantity — derived with Fer streamers in Theorem 2.1.4 for the uniform regimes (1.1.6)–(1.1.7).

One of the main differences is of course that while some of the constants in the big $\mathcal{O}$ notation in Theorem 2.1.1 grow with increasing λ, those in the big $\mathcal{O}$ notation in Theorem 2.1.4 can always be bounded independently of λ.

A striking manifestation of this difference between unbounded and bounded constants in the big $\mathcal{O}$ notation in, respectively, Theorem 2.1.1 and Theorem 2.1.4, is that while Theorem 2.1.1 and Theorem 2.1.4 share the same type of exponential growth in order, the rate is quite different. More concretely, the order with unbounded constants in the big $\mathcal{O}$ notation in Theorem 2.1.1 is

$$h_{\max}^{4\times 2^{l-1}-1},$$

whereas the order with bounded constants in the big $\mathcal{O}$ notation in Theorem 2.1.4 is

$$h_{\max}^{3\times 2^{l-1}-1}.$$

It is interesting to note that different manifestations of the same phenomenon have been reported also in the form of the number of function evaluations required to attain prescribed order, where moving from the asymptotic regimes (1.1.2)–(1.1.3) to the uniform regimes (1.1.6)–(1.1.7), leads to a localized increase in function evaluations for 'intermediary' eigenvalues (c.f., Subsection 1.1.2).

Definition 2.1.6. *Let $n \in \mathbb{Z}^+$, and define the*

$$\begin{aligned}
&\textit{exact flow:} && \boldsymbol{F}_\lambda(c_k, c_{k+1}) := \prod_{l=0}^{\infty} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})},\\
&\textit{exact solution:} && \boldsymbol{Y}_\lambda(c_{k+1}) = \boldsymbol{F}_\lambda(c_k, c_{k+1})\cdots\boldsymbol{F}_\lambda(c_1, c_2)\boldsymbol{F}_\lambda(a, c_1),\\
&\textit{truncated flow:} && \tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1}) := \prod_{l=0}^{n} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})},\\
&\textit{truncated solution:} && \tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1}) := \tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1})\cdots\tilde{\boldsymbol{F}}_{\lambda,n}(c_1, c_2)\tilde{\boldsymbol{F}}_{\lambda,n}(a, c_1),\\
&\textit{truncation local error:} && \boldsymbol{L}_{\lambda,n}^{trun.}(c_k, c_{k+1}) := \log\left(\boldsymbol{F}_\lambda(c_k, c_{k+1})\tilde{\boldsymbol{F}}_{\lambda,n}^{-1}(c_k, c_{k+1})\right),\\
&\textit{truncation global error:} && \boldsymbol{G}_{\lambda,n}^{trun.}(c_{k+1}) := \log\left(\boldsymbol{Y}_\lambda(c_{k+1})\tilde{\boldsymbol{Y}}_{\lambda,n}^{-1}(c_{k+1})\right).
\end{aligned}$$

Remark 2.1.7 (Ramos and Iserles, 2015)**.** *Observe that Remark 2.1.2 and Definition 2.1.6 ensure that the exact flow, the exact solution, the truncated flow and the truncated solution are in $\mathrm{SL}(2,\mathbb{R})$, and that the truncation local error and the truncation global error are in $\mathfrak{sl}(2,\mathbb{R})$. In particular, note that the truncated solution retains the geometric feature described in Subsection 1.1.4.*

Theorem 2.1.5 (Ramos and Iserles, 2015). *If Assumption 1.1.1 holds true, and $n \in \mathbb{Z}^+$, then, in the uniform regime (1.1.6), it follows that*

$$\boldsymbol{\pi}\left(\boldsymbol{L}_{\lambda,n}^{trun.}(c_k, c_{k+1})\right) = \delta_{|q'|}^{2^n} h_{\max}^{3\times 2^n - 1} \begin{bmatrix} \mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}_{\lambda,n}^{trun.}(c_{k+1})\right) = \delta_{|q'|}^{2^n} h_{\max}^{3\times 2^n - 2} \begin{bmatrix} \mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

and, in the uniform regime (1.1.7), it follows that

$$\boldsymbol{\pi}\left(\boldsymbol{L}_{\lambda,n}^{trun.}(c_k, c_{k+1})\right) = \delta_{|q'|}^{2^n} h_{\max}^{2^{n+1}} (\lambda - q_{\max})^{-\frac{2^n-1}{2}} \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}_{\lambda,n}^{trun.}(c_{k+1})\right) = \delta_{|q'|}^{2^n} h_{\max}^{2^{n+1}-1} (\lambda - q_{\max})^{-\frac{2^n-1}{2}} \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

where the constants in the big $\mathcal{O}$ notation can be bounded independently of λ.

Proof. See Section 2.5. □

Corollary 2.1.1 (Ramos and Iserles, 2015). *If Assumption 1.1.1 is true, and $n \in \mathbb{Z}^+$, then, in the two uniform regimes (1.1.6) and (1.1.7),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}_{\lambda,n}^{trun.}(c_k, c_{k+1})\right) = \delta_{|q'|}^{2^n} h_{\max}^{3\times 2^n - 1} \begin{bmatrix} \mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}_{\lambda,n}^{trun.}(c_{k+1})\right) = \delta_{|q'|}^{2^n} h_{\max}^{3\times 2^n - 2} \begin{bmatrix} \mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

where the constants in the big $\mathcal{O}$ notation can be bounded independently of λ.

Corollary 2.1.1 embodies the main result of this chapter, and it is worthwhile to pause and analyze its significance. As mentioned at the beginning of the current Subsection 2.1.2, the closed-forms provided by Fer streamers in Theorem 2.1.3, permit the development of a non-standard truncation of Fer expansions, with one less level of truncation, when compared with the two stage truncation process discussed at the end of the previous Subsection 2.1.1. The aforementioned non-standard truncation, relies on Fer streamers to sum up the infinite series in Fer expansions, which, in turn, lead to the approximation of the exact flow $\boldsymbol{F}_\lambda(c_k, c_{k+1})$ by the truncated flow $\tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1})$ and the approximation of the exact solution $\boldsymbol{Y}_\lambda(c_{k+1})$ by the truncated solution $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$, as specified in Definition 2.1.6. The results in Corollary 2.1.1 then assure that these approximations possess the advantageous properties listed at the start of the present chapter; indeed, in view of Assumption 1.1.1, it yields even more since the truncated approximations:

(i) do not impose any restriction on the step size for eigenvalues which are greater than or equal to the minimum of the potential,

(ii) require only a mild restriction on the step size for the remaining finite number of eigenvalues,

(iii) can attain any convergence rate, which grows exponentially with the number of terms, and are uniform for every eigenvalue in the sense of (1.1.6)–(1.1.7), and,

(iv) lend themselves to a clear understanding of the manner in which the potential affects the local and global truncation errors.

2.2 Conclusions

In view of Corollary 2.1.1, it is clear that the truncated solution $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$ with $n = 1, 2, 3, \ldots$ yields a uniform approximation of the exact solution $\boldsymbol{Y}_{\lambda}(c_{k+1})$, up to global order $4, 10, 22, \ldots$ uniformly with respect to (1.1.6) and (1.1.7).

Given Definition 2.1.6, this is an important result, since it reduces the problem of approximating the *infinite* product of exponentials in the exact flow $\boldsymbol{F}_{\lambda}(c_k, c_{k+1})$:

$$e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}, e^{\boldsymbol{D}_{\lambda,1}(c_k,c_{k+1})}, e^{\boldsymbol{D}_{\lambda,2}(c_k,c_{k+1})}, \ldots \tag{2.2.1}$$

to the problem of approximating the *finite* product of exponentials in the truncated flow $\tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1})$:

$$e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}, e^{\boldsymbol{D}_{\lambda,1}(c_k,c_{k+1})}, e^{\boldsymbol{D}_{\lambda,2}(c_k,c_{k+1})}, \ldots, e^{\boldsymbol{D}_{\lambda,n}(c_k,c_{k+1})}. \tag{2.2.2}$$

Going from theory to practice, to develop a numerical method, based on the error bounds of Corollary 2.1.1, that works equally well for all 'small', 'intermediary' and 'large' eigenvalues in the sense of (1.1.6)–(1.1.7), and, can attain arbitrary high-order, this means that we must devise a way to approximate (2.2.2), or, equivalently, given that the exponential map from $\mathfrak{sl}(2,\mathbb{R})$ to $\mathrm{SL}(2,\mathbb{R})$ has the simple closed-form expression in Definition 2.1.1, to approximate the finite number of exponents:

$$\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1}), \boldsymbol{D}_{\lambda,1}(c_k, c_{k+1}), \boldsymbol{D}_{\lambda,2}(c_k, c_{k+1}), \ldots, \boldsymbol{D}_{\lambda,n}(c_k, c_{k+1}), \tag{2.2.3}$$

the approximation of which will, in turn, give rise to a discretized flow $\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1})$ and to a discretized solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, which will themselves be computed exactly (c.f., Definition 3.4.2 in the next chapter).

Towards this end, it is important to note that, in light of Definition 2.1.2, each exponent in (2.2.3) is represented by a multivariate integral over a polytope, each of which is non-trivial, apart from the first. Indeed, even though the first exponent amounts to the

straightforward quadrature of the average of q in $[c_k, c_{k+1}]$:

$$\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1}) = \int_{c_k}^{c_{k+1}} \boldsymbol{B}_{\lambda,0}(c_k, t)dt = (c_{k+1} - c_k) \begin{bmatrix} 0 & 1 \\ \frac{\int_{c_k}^{c_{k+1}} q(\xi)d\xi}{c_{k+1}-c_k} - \lambda & 0 \end{bmatrix},$$

the remaining n exponents in (2.2.2) are far more complicated to approximate.

Fortunately, the Fer streamers closed-form in Theorem 2.1.3 yields an amenable expression to work with and again play a role of immense importance.

As an example, for the simplest non-trivial case, the intricacies associated with the quadrature of

$$\boldsymbol{D}_{\lambda,1}(c_k, c_{k+1}) = \int_{c_k}^{c_{k+1}} \boldsymbol{B}_{\lambda,1}(c_k, t)dt = h_k \int_0^1 \boldsymbol{B}_{\lambda,1}(c_k, c_k + h_k t)dt,$$

have already been touched upon in Remark 2.1.6 with the representation of the first Fer streamer in Remark 2.1.5. Namely, that $\boldsymbol{B}_{\lambda,1}(c_k, c_k + h_k \cdot) \in \mathfrak{sl}(2, \mathbb{R})$ is highly oscillatory whenever $\rho\left(\boldsymbol{D}_{\lambda,0}(c_k, c_k + h_k \cdot)\right) \in \mathbb{C}$ is purely imaginary with large norm. This observation is key since it is well-known that standard techniques such as Gauss–Christoffel quadrature are useless in the presence of highly oscillatory behaviour, and specialized techniques must be used instead.

Naturally, the remaining $n - 1$ exponents in (2.2.2) require even more care in the multivariate setting, where the challenges include:

- tracking down the behaviour of each multivariate integrand, by collecting like terms, to design each quadrature successfully,
- choosing the number of quadrature points of each quadrature conscientiously to be consistent with Corollary 2.1.1,
- choosing the set of quadrature points of each quadrature intelligently to minimize the error estimates whenever possible,
- decreasing the number of function evaluations to reduce the computational effort, and,
- decreasing the volume of linear algebra to reduce the computational effort.

In fact, the subject matter of Chapters 3, 4 and 5, lies precisely in the design, analysis and practical implementation of a discretized flow $\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$ (c.f., Definition 3.4.2) that are then used to develop the new numerical method of this dissertation, which, unlike previous approaches, is accompanied by error bounds that:

(i) hold uniformly over the entire eigenvalue range, in the sense of (1.1.6)–(1.1.7), and,

(ii) can attain arbitrary high-order.

2.3 Proof of Theorem 2.1.3

Note that

$$\begin{aligned}
\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right) &= \boldsymbol{\pi}\left(\sum_{j=1}^{\infty}(-1)^j\frac{j}{(j+1)!}\mathrm{ad}^j_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&= \sum_{j=1}^{\infty}(-1)^j\frac{j}{(j+1)!}\boldsymbol{\pi}\left(\mathrm{ad}^j_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&= \left(\sum_{j=1}^{\infty}(-1)^j\frac{j}{(j+1)!}\mathscr{C}^j_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\right)\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&= -\left(\sum_{j=1}^{\infty}\frac{2j-1}{(2j)!}\mathscr{C}^{2j-1}_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\right)\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&\quad+\left(\sum_{j=1}^{\infty}\frac{2j}{(2j+1)!}\mathscr{C}^{2j}_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\right)\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&= -\left(\sum_{j=1}^{\infty}\frac{2j-1}{(2j)!}\rho^{2j-2}\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&\quad+\left(\sum_{j=1}^{\infty}\frac{2j}{(2j+1)!}\rho^{2j-2}\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&= \varphi\left(\rho\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)\\
&\quad+\phi\left(\rho\left(\boldsymbol{D}_{\lambda,l-1}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,l-1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l-1}(c_k,t)\right)
\end{aligned}$$

where the first equality is due to Definition 2.1.2, and the third and penultimate equalities are due to Theorem 2.1.2.

2.4 Proof of Theorem 2.1.4

Recall Definitions 2.1.1 and 2.1.2 and note that

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) = 2(t-c_k)\sqrt{\frac{\int_{c_k}^{t}q(\xi)d\xi}{t-c_k}-\lambda}. \tag{2.4.1}$$

Note further that, (2.4.1) and assumptions (1.1.12) and (1.1.13) ensure

$$\lambda \in \left[q_{\max} - h_{\max}^{-2}, q_{\min}\right] \Rightarrow \rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) \in \left[0, 2h_{\max}\sqrt{q_{\max}-\lambda}\right] \subseteq [0,2] \tag{2.4.2}$$

$$\lambda \in [q_{\min}, q_{\max}] \Rightarrow \left|\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right| \leq 2h_{\max}\sqrt{q_{\max}-q_{\min}} \leq 2 \tag{2.4.3}$$

$$\lambda \in \left[q_{\max}, q_{\max} + h_{\max}^{-2}\right] \Rightarrow \rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) \in i\left[0, 2h_{\max}\sqrt{\lambda - q_{\min}}\right] \subseteq i\left[0, 2\sqrt{2}\right] \tag{2.4.4}$$

$$\lambda \in \left[q_{\max} + h_{\max}^{-2}, +\infty\right) \Rightarrow \rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) \in i\left[2\left(t-c_k\right)\sqrt{\lambda - q_{\max}}, +\infty\right) \tag{2.4.5}$$

which, together with Definition 2.1.4 and Remark 2.1.4, lead to the following estimates, in the two uniform regimes (1.1.6) and (1.1.7):

$$\left|\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\right| \leq 2, \qquad \text{w.r.t (1.1.6)}, \tag{2.4.6}$$

$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\right| \leq 1, \qquad \text{w.r.t (1.1.6)}, \tag{2.4.7}$$

$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right| \leq 2, \qquad \text{w.r.t (1.1.6)}, \tag{2.4.8}$$

$$\left|\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\right| \leq \frac{\left(t-c_k\right)^{-1}}{\sqrt{\lambda - q_{\max}}}, \qquad \text{w.r.t (1.1.7)}, \tag{2.4.9}$$

$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\right| \leq \frac{1}{2}\frac{\left(t-c_k\right)^{-2}}{\lambda - q_{\max}}, \qquad \text{w.r.t (1.1.7)}, \tag{2.4.10}$$

$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right| \leq 2, \qquad \text{w.r.t (1.1.7)}. \tag{2.4.11}$$

2.4.1 Estimating $\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)\cdots\exp\left(\boldsymbol{D}_{\lambda,0}(a,c_1)\right)$

Firstly, in the uniform regime (1.1.6), we have

$$\begin{aligned}
e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} &= \cosh\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \\
&\quad + \frac{\sinh\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}}{\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}}\begin{bmatrix}0 & c_{k+1}-c_k\\ \left(\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}\right)^2\left(c_{k+1}-c_k\right)^{-1} & 0\end{bmatrix} \\
&= \mathcal{O}(1)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \mathcal{O}(1)\begin{bmatrix}0 & \mathcal{O}(1)h_{\max}\\ \mathcal{O}(1)h_{\min}^{-1} & 0\end{bmatrix} \\
&= \mathcal{O}(1)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \mathcal{O}(1)\begin{bmatrix}0 & \mathcal{O}(1)h_{\max}\\ \mathcal{O}(1)h_{\max}^{-1} & 0\end{bmatrix}
\end{aligned}$$

where the first equality follows from Definition 2.1.1 and the second and third equalities follow from (1.1.14), (2.4.2), (2.4.3) and (2.4.4). Secondly, in the uniform regime (1.1.7),

we have

$$
\begin{aligned}
e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} &= \cos\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2i}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\\
&\quad+\sin\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2i}\begin{bmatrix}0 & \frac{c_{k+1}-c_k}{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\\ -\frac{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{c_{k+1}-c_k} & 0\end{bmatrix}\\
&= \mathcal{O}(1)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\mathcal{O}(1)\begin{bmatrix}0 & \frac{1}{\sqrt{\lambda-\frac{\int_{c_k}^{c_{k+1}}q(\xi)d\xi}{c_{k+1}-c_k}}}\\ -\sqrt{\lambda-\frac{\int_{c_k}^{c_{k+1}}q(\xi)d\xi}{c_{k+1}-c_k}} & 0\end{bmatrix}\\
&= \mathcal{O}(1)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\mathcal{O}(1)\begin{bmatrix}0 & \mathcal{O}(1)\frac{1}{\sqrt{\lambda-q_{\max}}}\\ \mathcal{O}(1)\sqrt{\lambda-q_{\min}} & 0\end{bmatrix}\\
&= \mathcal{O}(1)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\mathcal{O}(1)\begin{bmatrix}0 & \mathcal{O}(1)\frac{1}{\sqrt{\lambda-q_{\max}}}\\ \mathcal{O}(1)\sqrt{\lambda-q_{\max}} & 0\end{bmatrix}
\end{aligned}
$$

where the first equality follows from Definition 2.1.1, the second equality is due to (2.4.1) and (2.4.5), and the last equality is due to the fact that (1.1.12) ensures that

$$
\frac{\sqrt{\lambda-q_{\min}}}{\sqrt{\lambda-q_{\max}}}=\sqrt{1+\frac{q_{\max}-q_{\min}}{\lambda-q_{\max}}}\le\sqrt{1+h_{\max}^2(q_{\max}-q_{\min})}\le\sqrt{2}.
$$

2.4.2 Estimating $\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)$ and $\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)$

Finally, we note that (2.4.6)–(2.4.11), in turn, imply that

$$\begin{aligned}
&\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right)=\\
&=\begin{bmatrix}\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\frac{q(t)-\frac{\int_{c_k}^{t}q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)^2\\ 0\\ 0\end{bmatrix}\\
&=\begin{cases}\delta_{|q'|}\begin{bmatrix}\mathcal{O}\left(h_{\max}^2\right)\\ 0\\ 0\end{bmatrix}, & \text{w.r.t (1.1.6)},\\ \delta_{|q'|}\begin{bmatrix}\mathcal{O}\left(h_{\max}\right)\left(\lambda-q_{\max}\right)^{-\frac{1}{2}}\\ 0\\ 0\end{bmatrix}, & \text{w.r.t (1.1.7)},\end{cases}
\end{aligned}$$

and

$$\begin{aligned}
&\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right)=\\
&=\begin{bmatrix}0\\ -2\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\frac{q(t)-\frac{\int_{c_k}^{t}q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)^3\\ \frac{1}{2}\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\frac{q(t)-\frac{\int_{c_k}^{t}q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)\end{bmatrix}\\
&=\begin{cases}\delta_{|q'|}\begin{bmatrix}0\\ \mathcal{O}\left(h_{\max}^3\right)\\ \mathcal{O}\left(h_{\max}\right)\end{bmatrix}, & \text{w.r.t (1.1.6)},\\ \delta_{|q'|}\begin{bmatrix}0\\ \mathcal{O}\left(h_{\max}\right)\left(\lambda-q_{\max}\right)^{-1}\\ \mathcal{O}\left(h_{\max}\right)\end{bmatrix}, & \text{w.r.t (1.1.7)},\end{cases}
\end{aligned}$$

which, according to Theorem 2.1.3, lead to

$$\begin{aligned}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right) &= \varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right)\\ &\quad+\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right)\\ &= \begin{cases}\delta_{|q'|}\begin{bmatrix}\mathcal{O}\left(h_{\max}^2\right)\\ \mathcal{O}\left(h_{\max}^3\right)\\ \mathcal{O}\left(h_{\max}\right)\end{bmatrix}, & \text{w.r.t (1.1.6),}\\ \delta_{|q'|}\begin{bmatrix}\mathcal{O}\left(h_{\max}\right)\left(\lambda-q_{\max}\right)^{-\frac{1}{2}}\\ \mathcal{O}\left(h_{\max}\right)\left(\lambda-q_{\max}\right)^{-1}\\ \mathcal{O}\left(h_{\max}\right)\end{bmatrix}, & \text{w.r.t (1.1.7).}\end{cases}\end{aligned}$$

2.4.3 Estimating $\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right)$ and $\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)$ for $l \geq 2$

Follows by induction.

2.4.3.1 First step: $l = 2$

Given Definition 2.1.4 and the uniform estimates for $\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)$ in the previous subsection, it is now clear that

$$\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right) = \begin{cases}-\frac{1}{2}+\delta_{|q'|}^2\mathcal{O}\left(h_{\max}^6\right), & \text{w.r.t (1.1.6),}\\ -\frac{1}{2}+\delta_{|q'|}^2\mathcal{O}\left(h_{\max}^4\right)\left(\lambda-q_{\max}\right)^{-1}, & \text{w.r.t (1.1.7),}\end{cases}$$

$$\phi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right) = \begin{cases}\frac{1}{3}+\delta_{|q'|}^2\mathcal{O}\left(h_{\max}^6\right), & \text{w.r.t (1.1.6),}\\ \frac{1}{3}+\delta_{|q'|}^2\mathcal{O}\left(h_{\max}^4\right)\left(\lambda-q_{\max}\right)^{-1}, & \text{w.r.t (1.1.7),}\end{cases}$$

and, according to Theorem 2.1.3, that

$$\begin{aligned}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,2}(c_k,t)\right) &= \varphi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)\\ &\quad+\phi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)\\ &= \begin{cases}\delta^2_{|q'|}\begin{bmatrix}\mathcal{O}\left(h^5_{\max}\right)\\ \mathcal{O}\left(h^6_{\max}\right)\\ \mathcal{O}\left(h^4_{\max}\right)\end{bmatrix}, & \text{w.r.t (1.1.6)},\\ \delta^2_{|q'|}\begin{bmatrix}\mathcal{O}\left(h^3_{\max}\right)(\lambda-q_{\max})^{-1}\\ \mathcal{O}\left(h^3_{\max}\right)(\lambda-q_{\max})^{-\frac{3}{2}}\\ \mathcal{O}\left(h^3_{\max}\right)(\lambda-q_{\max})^{-\frac{1}{2}}\end{bmatrix}, & \text{w.r.t (1.1.7)}.\end{cases}\end{aligned}$$

2.4.3.2 Induction step: $l \Rightarrow l+1$

Given the induction claim, it is now clear that

$$\begin{aligned}\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right) &= \begin{cases}-\frac{1}{2}+\delta^{2^l}_{|q'|}\mathcal{O}\left(h^{3\times 2^l}_{\max}\right), & \text{w.r.t (1.1.6)},\\ -\frac{1}{2}+\delta^{2^l}_{|q'|}\mathcal{O}\left(h^{2^{l+1}}_{\max}\right)(\lambda-q_{\max})^{-2^{l-1}}, & \text{w.r.t (1.1.7)},\end{cases}\\ \phi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right) &= \begin{cases}\frac{1}{3}+\delta^{2^l}_{|q'|}\mathcal{O}\left(h^{3\times 2^l}_{\max}\right), & \text{w.r.t (1.1.6)},\\ \frac{1}{3}+\delta^{2^l}_{|q'|}\mathcal{O}\left(h^{2^{l+1}}_{\max}\right)(\lambda-q_{\max})^{-2^{l-1}}, & \text{w.r.t (1.1.7)},\end{cases}\end{aligned}$$

and, according to Theorem 2.1.3, that

$$\begin{aligned}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l+1}(c_k,t)\right) &= \varphi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,l}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right)\\ &\quad+\phi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,l}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right)\\ &= \begin{cases}\delta^{2^l}_{|q'|}\begin{bmatrix}\mathcal{O}\left(h^{3\times 2^l-1}_{\max}\right)\\ \mathcal{O}\left(h^{3\times 2^l}_{\max}\right)\\ \mathcal{O}\left(h^{3\times 2^l-2}_{\max}\right)\end{bmatrix}, & \text{w.r.t (1.1.6)},\\ \delta^{2^l}_{|q'|}\begin{bmatrix}\mathcal{O}\left(h^{2^{l+1}-1}_{\max}\right)(\lambda-q_{\max})^{-\frac{2^l}{2}}\\ \mathcal{O}\left(h^{2^{l+1}-1}_{\max}\right)(\lambda-q_{\max})^{-\frac{2^l+1}{2}}\\ \mathcal{O}\left(h^{2^{l+1}-1}_{\max}\right)(\lambda-q_{\max})^{-\frac{2^l-1}{2}}\end{bmatrix}, & \text{w.r.t (1.1.7)}.\end{cases}\end{aligned}$$

2.5 Proof of Theorem 2.1.5

The main obstacle in estimating the local and global truncation errors in Definition 2.1.6, resides in the fact that the lower-left entry of $\exp\left(\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1})\right)$ can be arbitrarily large, as testified by Theorem 2.1.4. This main obstacle can be circumvented by calling upon three Baker–Campbell–Hausdorff (BCH) type formulas

$$e^{\boldsymbol{X}}e^{\boldsymbol{Y}} = e^{\boldsymbol{X}+\boldsymbol{Y}+\frac{1}{2}[\boldsymbol{X},\boldsymbol{Y}]+\frac{1}{12}([\boldsymbol{X},[\boldsymbol{X},\boldsymbol{Y}]]+[\boldsymbol{Y},[\boldsymbol{Y},\boldsymbol{X}]])+\cdots} \tag{2.5.1}$$

$$e^{\boldsymbol{X}}e^{\boldsymbol{Y}}e^{-\boldsymbol{X}} = e^{\boldsymbol{Y}+[\boldsymbol{X},\boldsymbol{Y}]+\frac{1}{2}[\boldsymbol{X},[\boldsymbol{X},\boldsymbol{Y}]]+\frac{1}{6}[\boldsymbol{X},[\boldsymbol{X},[\boldsymbol{X},\boldsymbol{Y}]]]+\cdots} \tag{2.5.2}$$

$$= \exp\left(\mathrm{Ad}_{\exp(\boldsymbol{X})}\left(\boldsymbol{Y}\right)\right). \tag{2.5.3}$$

The truncation local error can be written as

$$\begin{aligned}
\boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k, c_{k+1}) &= \log\left(\boldsymbol{F}_\lambda(c_k, c_{k+1})\tilde{\boldsymbol{F}}^{-1}_{\lambda,n}(c_k, c_{k+1})\right) \\
&= \log\left(\left(\prod_{l=0}^{\infty} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)\left(\prod_{l=0}^{n} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)^{-1}\right) \\
&= \log\left(\left(\prod_{l=0}^{n} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)\left(\prod_{l=n+1}^{\infty} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)\left(\prod_{l=0}^{n} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)^{-1}\right) \\
&= \log\left(\left(\prod_{l=0}^{n} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right) e^{\boldsymbol{D}_{\lambda,n+1}(c_k,c_{k+1})+\text{h.o.t.}}\left(\prod_{l=0}^{n} e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)^{-1}\right) \\
&= \log\left(e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{D}_{\lambda,n+1}(c_k,c_{k+1})+\text{h.o.t.}}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\right) \\
&= \mathrm{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\left(\boldsymbol{D}_{\lambda,n+1}(c_k, c_{k+1}) + \text{h.o.t.}\right) \\
&= \mathrm{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\left(\boldsymbol{D}_{\lambda,n+1}(c_k, c_{k+1})\right) + \text{h.o.t.}
\end{aligned} \tag{2.5.4}$$

where the first and second equalities are due to Definition 2.1.6, the fourth equality is due to (2.5.1), the fifth equality is due to (2.5.2), and the sixth equality is due to (2.5.3). The local error expression (2.5.4), together with Theorem 2.1.4, yields the desired estimate.

The truncation global error obeys the recursion relation with initial condition

$$\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_1) = \boldsymbol{L}^{\text{trun.}}_{\lambda,n}(a, c_1) \tag{2.5.5}$$

and general rule

$$
\begin{aligned}
\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_{k+1}) &= \log\left(\boldsymbol{Y}_\lambda(c_{k+1})\tilde{\boldsymbol{Y}}^{-1}_{\lambda,n}(c_{k+1})\right)\\
&= \log\left(\boldsymbol{F}_\lambda(c_k,c_{k+1})\boldsymbol{Y}_\lambda(c_k)\tilde{\boldsymbol{Y}}^{-1}_{\lambda,n}(c_k)\tilde{\boldsymbol{F}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(\boldsymbol{F}_\lambda(c_k,c_{k+1})e^{\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_k)}\tilde{\boldsymbol{F}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k,c_{k+1})}\tilde{\boldsymbol{F}}_{\lambda,n}(c_k,c_{k+1})e^{\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_k)}\tilde{\boldsymbol{F}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k,c_{k+1})}\left(\prod_{l=0}^{n}e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)e^{\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_k)}\left(\prod_{l=0}^{n}e^{\boldsymbol{D}_{\lambda,l}(c_k,c_{k+1})}\right)^{-1}\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k,c_{k+1})}e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_k)+\text{h.o.t.}}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k,c_{k+1})}\exp\left(\text{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\left(\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_k)+\text{h.o.t.}\right)\right)\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k,c_{k+1})}\exp\left(\text{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\left(\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_k)\right)+\text{h.o.t.}\right)\right)\\
&= \boldsymbol{L}^{\text{trun.}}_{\lambda,n}(c_k,c_{k+1})+\text{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\left(\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_k)\right)+\text{h.o.t.}
\end{aligned}
\tag{2.5.6}
$$

where the first, second, third, fourth and fifth equalities are due to Definition 2.1.6, the sixth equality is due to (2.5.2), the seventh equality is due to (2.5.3), and the last equality is due to (2.5.1). The global error expressions (2.5.5) and (2.5.6) lead to

$$
\begin{aligned}
\boldsymbol{G}^{\text{trun.}}_{\lambda,n}(c_{k+1}) = {}& \text{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\left(\boldsymbol{D}_{\lambda,n+1}(c_k,c_{k+1})\right)\\
&+\text{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)\exp\left(\boldsymbol{D}_{\lambda,0}(c_{k-1},c_k)\right)}\left(\boldsymbol{D}_{\lambda,n+1}(c_{k-1},c_k)\right)\\
&+\cdots\\
&+\text{Ad}_{\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)\cdots\exp\left(\boldsymbol{D}_{\lambda,0}(a,c_1)\right)}\left(\boldsymbol{D}_{\lambda,n+1}(a,c_1)\right)\\
&+\text{h.o.t.}
\end{aligned}
$$

which, together with (1.1.14) and Theorem 2.1.4, result in the desired estimate.

Chapter 3

Retaining Fer streamers' properties under discretization

We have seen in Chapters 1–2 that, using a root-finding algorithm together with Theorems 1.0.1–1.0.2, one can, in principle, approximate uniformly all eigenvalues of the regular Sturm–Liouville problem (1.0.1)–(1.0.2), provided one can compute without any error the exact flow $\boldsymbol{F}_\lambda(c_k, c_{k+1})$ and solution $\boldsymbol{Y}_\lambda(c_{k+1})$ of the initial value problem (1.0.4)–(1.0.5) (c.f., Subsection 1.1.4 and Definition 2.1.6).

Of course, in view of Definition 2.1.6, this is in general not possible, since $\boldsymbol{F}_\lambda(c_k, c_{k+1})$ and $\boldsymbol{Y}_\lambda(c_{k+1})$ are each represented by an *infinite* product of exponentials (2.2.1), given by the Fer expansions.

However, we have also seen in Section 2.2 that via Fer streamers, i.e., via the closed-form expressions in Theorem 2.1.3 for each infinite sum in Definition 2.1.2, one can then establish Corollary 2.1.1 which, in essence, says that even though one cannot compute without any error the exact flow $\boldsymbol{F}_\lambda(c_k, c_{k+1})$ and solution $\boldsymbol{Y}_\lambda(c_{k+1})$, one can truncate each infinite product of exponentials while incurring only a small error, which is controlled by error bounds that:

(i) hold uniformly over the entire eigenvalue range, in the sense of (1.1.6)–(1.1.7), and,

(ii) can attain arbitrary high-order,

thereby giving rise to the truncated flow $\tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$, as bona fide approximations, each given by a *finite* product of exponentials (2.2.2), as prescribed by Definition 2.1.6.

Even though the uniform approximation of the exact flow $\boldsymbol{F}_\lambda(c_k, c_{k+1})$ and solution $\boldsymbol{Y}_\lambda(c_{k+1})$, by the truncated flow $\tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$, is a significant step, one should note however that, once again, the truncated flow $\tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$, cannot be computed without any error, simply because each exponential

in (2.2.2), or, equivalently, each exponent in (2.2.3), requires multivariate quadrature, and cannot be computed exactly, as touched upon in Section 2.2.

This motivates another approximation, that forms the subject of the present chapter, which reports on the work in (Ramos, 2015*a*). Namely, a uniform approximation this time of the truncated flow $\tilde{\boldsymbol{F}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$ by a discretized flow $\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, which are defined by the aforementioned multivariate quadrature of each exponent in (2.2.3), and hence can be computed exactly (c.f., Definition 3.4.2).

In the current chapter, we focus on developing such a discretized flow $\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, with uniform global orders less than or equal to 13. In light of the discussion in Section 2.2, we need to consider at most a uniform approximation via multivariate quadrature of

$$\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1}), \boldsymbol{D}_{\lambda,1}(c_k, c_{k+1}), \boldsymbol{D}_{\lambda,2}(c_k, c_{k+1}), \boldsymbol{D}_{\lambda,3}(c_k, c_{k+1}). \tag{3.0.1}$$

As mentioned above, it is the uniform and high-order advantageous features of the error bounds provided in Corollary 2.1.1 for the truncation error in the approximation of $\boldsymbol{Y}_\lambda(c_{k+1})$ by $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$, that forms the motivation to pursue the work in the current chapter to put forth and control the discretization error in the approximation of $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$ by $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, since it opens the door to the rigorous study over the entire eigenvalue range of the total error in the approximation of $\boldsymbol{Y}_\lambda(c_{k+1})$ by $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$.

The current chapter, pursues this rigorous study and shows that it is possible and affordable to discretize $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$ with $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, while retaining this uniformity in the total errors for all 'small', 'intermediary' and 'large' eigenvalues, with large step sizes uniform over the entire eigenvalue range. In particular, the discretization schemes are shown to enjoy large step sizes uniform over the entire eigenvalue range and tight error estimates uniform for every eigenvalue. They are made explicit for global orders $4, 7, 10$ and 13.

In addition, the present chapter provides total error estimates between $\boldsymbol{Y}_\lambda(c_{k+1})$ and $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, that quantify the interplay between the truncation and the discretization in the approach by Fer streamers.

This chapter is long and technical, since this is necessary in order to provide the Fer streamers approach with rigorous bounds on its total error while minimizing its number of function evaluations and volume of linear algebra. In particular, Section 3.1 derives the multivariate integrals for each exponent in (3.0.1), for global orders $4, 7, 10$ and 13, which occupy the central role in this chapter. Section 3.2 discusses the multivariate quadrature based on the simplest representation of each integrand, which makes its behaviour explicit — the first step towards any meaningful quadrature. This simplest representation is shown not to exploit the magnitude and behaviour of each integrand, and to lead to more function

evaluations and linear algebra than necessary. More importantly, the work in this section shows how to construct different representations that do exploit these features. Section 3.3 builds upon the previous section to put forth an alternative representation that leads to multivariate quadrature with less function evaluations and linear algebra. Section 3.4 quantifies the quadrature error in the former section and describes the total error estimates in the Fer streamers approach. Finally, Section 3.5 presents numerical results.

3.1 Multivariate integrals over polytopes

Following the discussion above, the present section puts forward the multivariate integrals required for each of the four terms in (3.0.1), for global orders $4, 7, 10$ and 13, which play the central role in this chapter.

As noted in Section 2.2, the first term in (3.0.1) amounts to the quadrature of

$$\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1}) = (c_{k+1} - c_k) \begin{bmatrix} 0 & 1 \\ \frac{\int_{c_k}^{c_{k+1}} q(\xi)d\xi}{c_{k+1}-c_k} - \lambda & 0 \end{bmatrix},$$

which we assume can be carried out without concern. In the current chapter it is important to note that the second term in (3.0.1) can be written as

$$\boldsymbol{D}_{\lambda,1}(c_k, c_{k+1}) = \int_{c_k}^{c_{k+1}} \boldsymbol{B}_{\lambda,1}(c_k, t)dt,$$

the third term in (3.0.1) can be controlled by

$$\begin{aligned}
&\boldsymbol{D}_{\lambda,2}(c_k, c_{k+1}) \\
&= \int_{c_k}^{c_{k+1}} \boldsymbol{B}_{\lambda,2}(c_k, t)dt \\
&= \int_{c_k}^{c_{k+1}} \varphi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k, t)\right)\right) \mathrm{ad}_{\boldsymbol{D}_{\lambda,1}(c_k,t)} \boldsymbol{B}_{\lambda,1}(c_k, t)dt \\
&\quad + \int_{c_k}^{c_{k+1}} \phi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k, t)\right)\right) \mathrm{ad}^2_{\boldsymbol{D}_{\lambda,1}(c_k,t)} \boldsymbol{B}_{\lambda,1}(c_k, t)dt
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{2}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)]\,dt_2dt_1\\
&\quad+\frac{1}{3}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}[\boldsymbol{B}_{\lambda,1}(c_k,t_3),[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)]]\,dt_3dt_2dt_1\\
&\quad-\frac{1}{8}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\int_{c_k}^{t_1}[\boldsymbol{B}_{\lambda,1}(c_k,t_4),[\boldsymbol{B}_{\lambda,1}(c_k,t_3),[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)]]]\,dt_4dt_3dt_2dt_1\\
&\quad+\begin{cases}\delta^5_{|q'|}h^{14}_{\max}\boldsymbol{\pi}^{-1}\left(\begin{bmatrix}\mathcal{O}(h_{\max}) & \mathcal{O}(h^2_{\max}) & \mathcal{O}(1)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.6)},\\ \delta^5_{|q'|}h^{10}_{\max}(\lambda-q_{\max})^{-2}\boldsymbol{\pi}^{-1}\left(\begin{bmatrix}\frac{\mathcal{O}(1)}{\sqrt{\lambda-q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda-q_{\max}} & \mathcal{O}(1)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.7)},\end{cases}\\
&= -\frac{1}{2}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)]\,dt_2dt_1\\
&\quad+\frac{1}{3}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}[\boldsymbol{B}_{\lambda,1}(c_k,t_3),[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)]]\,dt_3dt_2dt_1\\
&\quad+\begin{cases}\delta^4_{|q'|}h^{11}_{\max}\boldsymbol{\pi}^{-1}\left(\begin{bmatrix}\mathcal{O}(h_{\max}) & \mathcal{O}(h^2_{\max}) & \mathcal{O}(1)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.6)},\\ \delta^4_{|q'|}h^{8}_{\max}(\lambda-q_{\max})^{-\frac{3}{2}}\boldsymbol{\pi}^{-1}\left(\begin{bmatrix}\frac{\mathcal{O}(1)}{\sqrt{\lambda-q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda-q_{\max}} & \mathcal{O}(1)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.7)},\end{cases}\\
&= -\frac{1}{2}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)]\,dt_2dt_1\\
&\quad+\begin{cases}\delta^3_{|q'|}h^{8}_{\max}\boldsymbol{\pi}^{-1}\left(\begin{bmatrix}\mathcal{O}(h_{\max}) & \mathcal{O}(h^2_{\max}) & \mathcal{O}(1)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.6)},\\ \delta^3_{|q'|}h^{6}_{\max}(\lambda-q_{\max})^{-1}\left(\begin{bmatrix}\frac{\mathcal{O}(1)}{\sqrt{\lambda-q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda-q_{\max}} & \mathcal{O}(1)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.7)},\end{cases}
\end{aligned}
$$

where the first equality follows from Definition 2.1.2, the second equality is due to Theorem 2.1.3 and the third, fourth and fifth equalities are due to Definition 2.1.2 and Theorem 2.1.4, and, similarly, the fourth term in (3.0.1) can be controlled by

$$
\begin{aligned}
&\boldsymbol{D}_{\lambda,3}(c_k,c_{k+1})\\
&=\int_{c_k}^{c_{k+1}}\boldsymbol{B}_{\lambda,3}(c_k,t)dt\\
&=\int_{c_k}^{c_{k+1}}\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,2}(c_k,t)\right)\right)\mathrm{ad}_{\boldsymbol{D}_{\lambda,2}(c_k,t)}\boldsymbol{B}_{\lambda,2}(c_k,t)dt\\
&\quad+\int_{c_k}^{c_{k+1}}\phi\left(\rho\left(\boldsymbol{D}_{\lambda,2}(c_k,t)\right)\right)\mathrm{ad}^2_{\boldsymbol{D}_{\lambda,2}(c_k,t)}\boldsymbol{B}_{\lambda,2}(c_k,t)dt
\end{aligned}
$$

$$= -\frac{1}{8}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\int_{c_k}^{t_2}\left[\left[\boldsymbol{B}_{\lambda,1}(c_k,t_4),\boldsymbol{B}_{\lambda,1}(c_k,t_2)\right],\left[\boldsymbol{B}_{\lambda,1}(c_k,t_3),\boldsymbol{B}_{\lambda,1}(c_k,t_1)\right]\right]dt_4dt_3dt_2dt_1$$
$$+\begin{cases}\delta_{|q'|}^5 h_{\max}^{14}\boldsymbol{\pi}^{-1}\left(\begin{bmatrix}\mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.6)},\\ \delta_{|q'|}^5 h_{\max}^{10}\left(\lambda-q_{\max}\right)^{-2}\boldsymbol{\pi}^{-1}\left(\begin{bmatrix}\frac{\mathcal{O}(1)}{\sqrt{\lambda-q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda-q_{\max}} & \mathcal{O}\left(1\right)\end{bmatrix}^\top\right), & \text{w.r.t (1.1.7)}.\end{cases}$$

In particular, the computation of the second, third and fourth terms in (3.0.1) boils down to the quadrature of the following multivariate integrals over polytopes:

$$\int_{c_k}^{c_{k+1}}\boldsymbol{B}_{\lambda,1}(c_k,t)dt, \tag{3.1.1}$$
$$-\frac{1}{2}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\left[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)\right]dt_2dt_1, \tag{3.1.2}$$
$$\frac{1}{3}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\left[\boldsymbol{B}_{\lambda,1}(c_k,t_3),\left[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)\right]\right]dt_3dt_2dt_1, \tag{3.1.3}$$
$$-\frac{1}{8}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\left[\boldsymbol{B}_{\lambda,1}(c_k,t_4),\left[\boldsymbol{B}_{\lambda,1}(c_k,t_3),\left[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)\right]\right]\right]dt_4dt_3dt_2dt_1, \tag{3.1.4}$$
$$-\frac{1}{8}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\int_{c_k}^{t_2}\left[\left[\boldsymbol{B}_{\lambda,1}(c_k,t_4),\boldsymbol{B}_{\lambda,1}(c_k,t_2)\right],\left[\boldsymbol{B}_{\lambda,1}(c_k,t_3),\boldsymbol{B}_{\lambda,1}(c_k,t_1)\right]\right]dt_4dt_3dt_2dt_1. \tag{3.1.5}$$

In detail, according to Definition 2.1.6 as well as to the equalities and estimates in the current section, in order to design Fer streamers with global order 4, 7, 10 and 13, a uniform approximation of the truncated flow $\tilde{\boldsymbol{F}}_{\lambda,n}(c_k,c_{k+1})$ and solution $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$, requires the quadrature of (3.1.1), (3.1.1)–(3.1.2), (3.1.1)–(3.1.3) and (3.1.1)–(3.1.5) up to local order 5, 8, 11 and 14, respectively, in the sense of the uniform regimes (1.1.6)–(1.1.7).

3.2 Towards an optimal quadrature

The present section discusses the multivariate quadrature of (3.1.1)–(3.1.5) based on the simplest representation of each integrand, which makes its behaviour explicit: the first requisite to develop any sensible quadrature. In particular, in line with the two cases distinguished in (1.1.12) and (1.1.13), these quadrature schemes are based on the partition of the eigenvalue interval $[q_{\max}-h_{\max}^{-2},+\infty)$ into three intervals:

$$\lambda\in\left[q_{\max}-h_{\max}^{-2},q_{\min}-1\right], \tag{3.2.1}$$
$$\lambda\in\left[q_{\min}-1,q_{\max}+1\right], \tag{3.2.2}$$
$$\lambda\in\left[q_{\max}+1,+\infty\right), \tag{3.2.3}$$

and are shown to possess the following advantages and disadvantages:

- Advantages:
  - They respect the behaviour of each integrand, and,
  - They are less technical than those in Section 3.3.
- Disadvantages:
  - They do not exploit the magnitude of each integrand as a means to reduce the number of function evaluations and volume of linear algebra, and,
  - They do not exploit the behaviour of each integrand as a means to decrease the quadrature error without using derivatives of the potential.

3.2.1 Representations with complex trigonometric polynomials

It is the aim of this subsection to construct representations of $\boldsymbol{B}_{\lambda,1}(c_k,t)$ which make its behaviour explicit. This is achieved in Theorems 3.2.1, 3.2.2 and 3.2.3 below.

To have intuition before going into details, the reader should be aware that the idea that leads to the representations in Theorems 3.2.1, 3.2.2 and 3.2.3 below is to rewrite the representation of $\boldsymbol{B}_{\lambda,1}(c_k,t)$ in Remark 2.1.5 in terms of exponential functions with argument $2(t-c_k)\sqrt{q(c_k)-\lambda}$. To this end, recall Remark 2.1.5 and call upon Definition 2.1.4 to rewrite

$$\varphi(z) = -\frac{1}{z^2} + \frac{1}{2}\left(\frac{1}{z^2} - \frac{1}{z}\right)e^{z} + \frac{1}{2}\left(\frac{1}{z^2} + \frac{1}{z}\right)e^{-z}, \tag{3.2.4}$$

$$\phi(z) = \frac{1}{2}\left(\frac{1}{z^2} - \frac{1}{z^3}\right)e^{z} + \frac{1}{2}\left(\frac{1}{z^2} + \frac{1}{z^3}\right)e^{-z}, \tag{3.2.5}$$

$$\phi(z)z^2 = \frac{1}{2}\left(1 - \frac{1}{z}\right)e^{z} + \frac{1}{2}\left(1 + \frac{1}{z}\right)e^{-z}, \tag{3.2.6}$$

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) = 2(t-c_k)\sqrt{\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k} - \lambda}, \tag{3.2.7}$$

$$e^{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)} = e^{2(t-c_k)\left(\sqrt{\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda}-\sqrt{q(c_k)-\lambda}\right)}e^{2(t-c_k)\sqrt{q(c_k)-\lambda}}. \tag{3.2.8}$$

Since

$$e^{2(t-c_k)\left(\sqrt{\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda}-\sqrt{q(c_k)-\lambda}\right)} \tag{3.2.9}$$

is close to 1 uniformly for every eigenvalue (c.f., Sections 3.7 and 3.8), the behaviour of $\boldsymbol{B}_{\lambda,1}(c_k,t)$ will be determined in terms of exponential functions with the argument $2(t-c_k)\sqrt{q(c_k)-\lambda}$. As will become clear, this will serve to make the behaviour of $\boldsymbol{B}_{\lambda,1}(c_k,t)$ explicit.

To make this idea precise, the following definition introduces the non-exponential parts $\zeta_{\lambda,1}(c_k,t) \in \mathbb{R}$, $\boldsymbol{R}_1 \in \mathfrak{sl}(2,\mathbb{R})$, $\boldsymbol{S}_{\lambda,1}(c_k,t) \in \mathfrak{sl}(2,\mathbb{C})$, $\boldsymbol{U}_{\lambda,1}(c_k,t) \in \mathfrak{sl}(2,\mathbb{R})$ and $\boldsymbol{V}_{\lambda,1}(c_k,t) \in \mathfrak{sl}(2,\mathbb{R})$ which appear below in Theorems 3.2.1 and 3.2.3. Although important, it is technical in nature and the reader is encouraged to glance over it and return to it as required.

Definition 3.2.1.

$$\zeta_{\lambda,1}(c_k,t) := \frac{1}{4}\frac{q(t)-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}{t-c_k}\frac{1}{\lambda-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}},$$

$$\zeta_{\lambda,1}(c_k,c_k) := \frac{q'\left(c_k^+\right)}{8\left(\lambda-q\left(c_k\right)\right)},$$

$$\boldsymbol{\pi}\left(\boldsymbol{R}_1\right) := \begin{bmatrix}1 & 0 & 0\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{S}_{\lambda,1}(c_k,t)\right) := \frac{1}{8}\frac{q(t)-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}{t-c_k}e^{2i(t-c_k)\left(\sqrt{\lambda-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}-\sqrt{\lambda-q(c_k)}\right)}$$
$$\times\begin{bmatrix}-\frac{1}{\lambda-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}+i\frac{2(t-c_k)}{\left(\lambda-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}\right)^{\frac{1}{2}}}\\ \frac{2(t-c_k)}{\lambda-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}+i\frac{1}{\left(\lambda-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}\right)^{\frac{3}{2}}}\\ 2\left(t-c_k\right)+i\frac{1}{\left(\lambda-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}\right)^{\frac{1}{2}}}\end{bmatrix},$$

$$\boldsymbol{\pi}\left(\boldsymbol{S}_{\lambda,1}(c_k,c_k)\right) := \frac{q'\left(c_k^+\right)}{16}\begin{bmatrix}-\frac{1}{\lambda-q(c_k)} & i\frac{1}{(\lambda-q(c_k))^{\frac{3}{2}}} & i\frac{1}{(\lambda-q(c_k))^{\frac{1}{2}}}\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{U}_{\lambda,1}(c_k,t)\right) := \frac{1}{8}\frac{q(t)-\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}}{t-c_k}e^{2(t-c_k)\left(\sqrt{\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda}-\sqrt{q(c_k)-\lambda}\right)}$$
$$\times\begin{bmatrix}\frac{1}{\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda}-\frac{2(t-c_k)}{\left(\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda\right)^{\frac{1}{2}}}\\ -\frac{2(t-c_k)}{\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda}+\frac{1}{\left(\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda\right)^{\frac{3}{2}}}\\ 2\left(t-c_k\right)-\frac{1}{\left(\frac{\int_{c_k}^t q(\xi)d\xi}{t-c_k}-\lambda\right)^{\frac{1}{2}}}\end{bmatrix},$$

$$\boldsymbol{\pi}\left(\boldsymbol{U}_{\lambda,1}(c_k,c_k)\right) := \frac{q'\left(c_k^+\right)}{16}\begin{bmatrix}\frac{1}{q(c_k)-\lambda} & \frac{1}{(q(c_k)-\lambda)^{\frac{3}{2}}} & -\frac{1}{(q(c_k)-\lambda)^{\frac{1}{2}}}\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{V}_{\lambda,1}(c_k,t)\right) := \frac{1}{8}\frac{q(t)-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}{t-c_k} e^{-2(t-c_k)\left(\sqrt{\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda}-\sqrt{q(c_k)-\lambda}\right)} \times \begin{bmatrix} \frac{1}{\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda} + \frac{2(t-c_k)}{\left(\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda\right)^{\frac{1}{2}}} \\ -\frac{2(t-c_k)}{\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda} - \frac{1}{\left(\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda\right)^{\frac{3}{2}}} \\ 2\left(t-c_k\right) + \frac{1}{\left(\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}-\lambda\right)^{\frac{1}{2}}} \end{bmatrix},$$

$$\boldsymbol{\pi}\left(\boldsymbol{V}_{\lambda,1}(c_k,c_k)\right) := \frac{q'\left(c_k^{+}\right)}{16}\begin{bmatrix} \frac{1}{q(c_k)-\lambda} & -\frac{1}{(q(c_k)-\lambda)^{\frac{3}{2}}} & \frac{1}{(q(c_k)-\lambda)^{\frac{1}{2}}} \end{bmatrix}^{\top}.$$

Focusing first on the left interval (3.2.1), we note that a closer investigation reveals that:

Theorem 3.2.1 (Ramos, 2015*a*). *If λ lies on (3.2.1) then*

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right) = \zeta_{\lambda,1}(c_k,t)\boldsymbol{\pi}\left(\boldsymbol{R}_1\right) + \boldsymbol{\pi}\left(\boldsymbol{U}_{\lambda,1}(c_k,t)\right) e^{2\sqrt{q(c_k)-\lambda}(t-c_k)} + \boldsymbol{\pi}\left(\boldsymbol{V}_{\lambda,1}(c_k,t)\right) e^{-2\sqrt{q(c_k)-\lambda}(t-c_k)}$$

where $\zeta_{\lambda,1}(c_k,t)$, $\boldsymbol{R}_1$, $\boldsymbol{U}_{\lambda,1}(c_k,t)$ and $\boldsymbol{V}_{\lambda,1}(c_k,t)$ are as in Definition 3.2.1. In addition, it is true that the derivatives $\zeta_{\lambda,1}^{(j)}(c_k,t)$, $\boldsymbol{U}_{\lambda,1}^{(j)}(c_k,t)$ and $\boldsymbol{V}_{\lambda,1}^{(j)}(c_k,t)$ can be bounded independently of λ.

Proof. See Section 3.7. □

Looking next at the middle interval (3.2.2), we rewrite:

Theorem 3.2.2 (Ramos, 2015*a*). *It is true that*

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right) = \begin{bmatrix} \varphi\left(\sqrt{\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)}\right)\frac{q(t)-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)^2 \\ -2\phi\left(\sqrt{\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)}\right)\frac{q(t)-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k)^3 \\ \frac{1}{2}\phi\left(\sqrt{\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)}\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\frac{q(t)-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}{t-c_k}(t-c_k) \end{bmatrix},$$

where (recall Definition 2.1.4)

$$\varphi(\sqrt{z}) := -\sum_{j=0}^{\infty} \frac{2j+1}{(2j+2)!} z^j,$$
$$\phi(\sqrt{z}) := \sum_{j=0}^{\infty} \frac{2j+2}{(2j+3)!} z^j,$$

are analytic in $z \in \mathbb{C}$*, and,*

$$\rho^2 \left(\boldsymbol{D}_{\lambda,0}(c_k, t) \right) = 4(t-c_k)^2 \left(\frac{\int_{c_k}^{t} q(\xi) d\xi}{t-c_k} - \lambda \right).$$

Furthermore, if λ *lies on* (3.2.2) *then the derivatives of* $\boldsymbol{B}_{\lambda,1}(c_k, t)$ *can be bounded independently of* λ*.*

Proof. See Remark 2.1.5. □

Looking next at the right interval (3.2.3), in line with Remark 2.1.6, we expose the oscillatory behaviour:

Theorem 3.2.3 (Ramos, 2015*a*)**.** *If* λ *lies on* (3.2.3) *then*

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right) = \zeta_{\lambda,1}(c_k,t)\boldsymbol{\pi}\left(\boldsymbol{R}_1\right) + \boldsymbol{\pi}\left(\boldsymbol{S}_{\lambda,1}(c_k,t)\right) e^{2i\sqrt{\lambda-q(c_k)}(t-c_k)} + \overline{\boldsymbol{\pi}\left(\boldsymbol{S}_{\lambda,1}(c_k,t)\right)} e^{-2i\sqrt{\lambda-q(c_k)}(t-c_k)}$$

where $\zeta_{\lambda,1}(c_k, t)$*,* $\boldsymbol{R}_1$ *and* $\boldsymbol{S}_{\lambda,1}(c_k, t)$ *are as in Definition 3.2.1. In addition, it is true that the derivatives* $\zeta^{(j)}_{\lambda,1}(c_k, t)$ *and* $\boldsymbol{S}^{(j)}_{\lambda,1}(c_k, t)$ *can be bounded independently of* λ*.*

Proof. See Section 3.8. □

3.2.2 Drawbacks with complex trigonometric polynomials

Given Theorems 3.2.1 and 3.2.3, for $\lambda \in \left[q_{\max} - h_{\max}^{-2}, q_{\min} - 1\right] \cup \left[q_{\max} + 1, +\infty\right)$, the construction of a quadrature which respects the behaviour of each integrand is possible by polynomial interpolation of $\zeta_{\lambda,1}(c_k, t)$, $\boldsymbol{S}_{\lambda,1}(c_k, t)$, $\boldsymbol{U}_{\lambda,1}(c_k, t)$ and $\boldsymbol{V}_{\lambda,1}(c_k, t)$ in $t \in [c_k, c_{k+1}]$ and the exact integration of the result. Similarly, given Theorem 3.2.2, for $\lambda \in \left[q_{\min} - 1, q_{\max} + 1\right]$, the construction of a quadrature is possible by polynomial interpolation of $\boldsymbol{B}_{\lambda,1}(c_k, t)$ in $t \in [c_k, c_{k+1}]$ and the exact integration of the result. These are Filon-type quadrature schemes.

The idea to turn multivariate quadrature into univariate polynomial interpolation was introduced in (Iserles and Nørsett, 1999*a*) for well-behaved integrands in the context of Magnus and Fer expansions and in (Iserles, 2004*b*) for highly oscillatory Fourier-type integrands in the context of modified Magnus expansions. In this chapter it is introduced

for a plethora of behaviours (mildly to highly oscillatory Fourier-type in Theorem 3.2.3, mildly to highly oscillatory unconventional-type in Theorem 3.3.3, etc.) in the context of Fer streamers.

Unfortunately, for $\lambda \in \left[q_{\max} - h_{\max}^{-2}, q_{\min} - 1\right] \cup \left[q_{\max} + 1, +\infty\right)$, these quadrature schemes present two major drawbacks, which can be traced back to the representations in Theorems 3.2.1 and 3.2.3.

The first drawback is particularly severe for $\lambda \in \left[q_{\max} + 1, +\infty\right)$ because of the mildly to highly oscillatory Fourier-type behaviour identified in Theorem 3.2.3. To pinpoint the issue, it is important to recall that, in recent years, it has been made clear with the theoretical analysis in (Levin, 1996; Iserles and Nørsett, 2005; Iserles and Nørsett, 2006), that the polynomial interpolation in Filon-type quadrature schemes with highly oscillatory Fourier-type behaviour should include the endpoints. As discussed in those papers, this is done because it makes the difference between the function and the interpolation polynomial equal to zero at the endpoints, which, in many cases, is shown to result in a decrease in quadrature error. With the representation of $\boldsymbol{B}_{\lambda,1}(c_k, t)$ in Theorem 3.2.3 in mind, this leads to the evaluation of $\boldsymbol{S}_{\lambda,1}(c_k, c_k)$ which depends on $q'(c_k^+)$ (c.f., Definition 3.2.1). Hence, it is not possible to exploit the representation of $\boldsymbol{B}_{\lambda,1}(c_k, t)$ in Theorem 3.2.3 as a means to decrease the quadrature error without using derivatives of the potential, which is not desirable from a computational point of view since the derivative of the potential might not be available in closed-form. A similar issue occurs also for the subset $\lambda \in \left[q_{\max} - h_{\max}^{-2}, q_{\min} - 1\right]$ given the behaviour exposed in Theorem 3.2.1 and the fact that $\boldsymbol{U}_{\lambda,1}(c_k, c_k)$ and $\boldsymbol{V}_{\lambda,1}(c_k, c_k)$ depend on $q'(c_k^+)$ (c.f., Definition 3.2.1).

The second drawback is equally acute for every $\lambda \in \left[q_{\max} - h_{\max}^{-2}, q_{\min} - 1\right]$ as well as $\lambda \in \left[q_{\max} + 1, +\infty\right)$. In short, Theorems 3.2.1 and 3.2.3 do not respect the magnitude of $\boldsymbol{B}_{\lambda,1}(c_k, t)$, i.e., they represent $\boldsymbol{B}_{\lambda,1}(c_k, t)$ in terms of larger quantities, which is not desirable from a computational point of view because it leads to more function evaluations and linear algebra. As an example, these results represent the zero vector

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k, c_k)\right) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^\top,$$

firstly in Theorem 3.2.1 as the sum of

$$-\frac{q'\left(c_k^+\right)}{8\left(q\left(c_k\right) - \lambda\right)} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \frac{q'\left(c_k^+\right)}{16} \begin{bmatrix} \frac{1}{q(c_k)-\lambda} \\ \frac{1}{(q(c_k)-\lambda)^{\frac{3}{2}}} \\ -\frac{1}{(q(c_k)-\lambda)^{\frac{1}{2}}} \end{bmatrix}, \qquad \frac{q'\left(c_k^+\right)}{16} \begin{bmatrix} \frac{1}{q(c_k)-\lambda} \\ -\frac{1}{(q(c_k)-\lambda)^{\frac{3}{2}}} \\ \frac{1}{(q(c_k)-\lambda)^{\frac{1}{2}}} \end{bmatrix},$$

and secondly in Theorem 3.2.3 as the sum of

$$\frac{q'\left(c_k^+\right)}{8\left(\lambda - q\left(c_k\right)\right)}\begin{bmatrix}1\\0\\0\end{bmatrix}, \qquad \frac{q'\left(c_k^+\right)}{16}\begin{bmatrix}-\frac{1}{\lambda-q(c_k)}\\ i\frac{1}{(\lambda-q(c_k))^{\frac{3}{2}}}\\ i\frac{1}{(\lambda-q(c_k))^{\frac{1}{2}}}\end{bmatrix}, \qquad \overline{\frac{q'\left(c_k^+\right)}{16}\begin{bmatrix}-\frac{1}{\lambda-q(c_k)}\\ i\frac{1}{(\lambda-q(c_k))^{\frac{3}{2}}}\\ i\frac{1}{(\lambda-q(c_k))^{\frac{1}{2}}}\end{bmatrix}}.$$

These are two unfortunate features of the representations of $\boldsymbol{B}_{\lambda,1}(c_k,t)$ in Theorems 3.2.1 and 3.2.3, and are not intrinsic properties of $\boldsymbol{B}_{\lambda,1}(c_k,t)$. In fact, in the next section, we present two different representations which, although more technical in nature, do not suffer from these issues.

3.3 Optimal quadrature

Given the intrinsic shortcomings of the quadrature schemes in the previous section, it is natural to ask whether these can be circumvented. This is the aim of the present section, which concerns quadrature schemes built on different representations of each integrand of the multivariate integrals (3.1.1)–(3.1.5), based on the same two cases distinguished in (1.1.12)–(1.1.13), but on the different partition of the eigenvalue interval $[q_{\max}-h_{\max}^{-2},+\infty)$ into three subsets:

$$\lambda \in [q_{\max} - h_{\max}^{-2}, q_{\min} - 1] \cup [q_{\max} + 1, q_{\max} + h_{\max}^{-2}], \tag{3.3.1}$$

$$\lambda \in [q_{\min} - 1, q_{\max} + 1], \tag{3.3.2}$$

$$\lambda \in [q_{\max} + h_{\max}^{-2}, +\infty), \tag{3.3.3}$$

with the following advantages and disadvantages:

- Advantages:
  - They respect the behaviour of each integrand,
  - They exploit the magnitude of each integrand as a means to reduce the number of function evaluations and volume of linear algebra, and,
  - They exploit the behaviour of each integrand as a means to decrease the quadrature error without using derivatives of the potential.
- Disadvantages:
  - They are more technical in nature than the ones in Section 3.2.

3.3.1 Representations with real trigonometric polynomials

It is the purpose of this subsection to develop representations of $\boldsymbol{B}_{\lambda,1}(c_k,t)$ which make both its behaviour and magnitude explicit. This is accomplished in Theorems 3.3.1, 3.3.2 and 3.3.3 below and exploited in the subsubsections that follow.

Definition 3.3.1. *Let*

$$\omega_{\lambda,1}(c_k,t) := 2(t-c_k)\sqrt{\lambda - q(c_k)}, \tag{3.3.4}$$

$$r_{\lambda,1}(c_k,t) := \sqrt{\frac{\lambda - \frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}{\lambda - q(c_k)}}, \tag{3.3.5}$$

$$\epsilon_{\lambda,1}(c_k,t) := \omega_{\lambda,1}(c_k,t)\left(r_{\lambda,1}(c_k,t)-1\right), \tag{3.3.6}$$

$$s_{\lambda,1}(c_k,t) := \omega_{\lambda,1}(c_k,t)\epsilon_{\lambda,1}(c_k,t). \tag{3.3.7}$$

To provide intuition before plunging into technicalities, the reader should be aware that the guiding principle that leads to the representations in Theorems 3.3.1, 3.3.2 and 3.3.3 below is to rewrite the representation of $\boldsymbol{B}_{\lambda,1}(c_k,t)$ in Remark 2.1.5 in terms of trigonometric functions with the argument $\omega_{\lambda,1}(c_k,t)$. To this end, recall Remark 2.1.5 and invoke Definitions 2.1.4 and 3.3.1 to rewrite

$$\varphi(z) = \frac{\cosh(z)-1}{z^2} - \frac{\sinh(z)}{z}, \tag{3.3.8}$$

$$\phi(z) = \frac{1}{z^2}\left(\cosh(z) - \frac{\sinh(z)}{z}\right), \tag{3.3.9}$$

$$\phi(z)z^2 = \cosh(z) - \frac{\sinh(z)}{z}, \tag{3.3.10}$$

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) = i\cdot 2(t-c_k)\sqrt{\lambda - \frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}} = i\cdot\omega_{\lambda,1}(c_k,t)r_{\lambda,1}(c_k,t), \tag{3.3.11}$$

$$\begin{aligned}\cosh\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right) &= \cos\left(\omega_{\lambda,1}(c_k,t)r_{\lambda,1}(c_k,t)\right)\\ &= \cos\left(\epsilon_{\lambda,1}(c_k,t)+\omega_{\lambda,1}(c_k,t)\right)\\ &= \cos\left(\epsilon_{\lambda,1}(c_k,t)\right)\cdot\cos\left(\omega_{\lambda,1}(c_k,t)\right)\\ &\quad - \sin\left(\epsilon_{\lambda,1}(c_k,t)\right)\cdot\sin\left(\omega_{\lambda,1}(c_k,t)\right),\end{aligned} \tag{3.3.12}$$

$$\begin{aligned}\frac{\sinh\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)}{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)} &= \frac{\sin\left(\omega_{\lambda,1}(c_k,t)r_{\lambda,1}(c_k,t)\right)}{\omega_{\lambda,1}(c_k,t)r_{\lambda,1}(c_k,t)}\\ &= \frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)+\omega_{\lambda,1}(c_k,t)\right)}{\omega_{\lambda,1}(c_k,t)r_{\lambda,1}(c_k,t)}\\ &= \frac{r_{\lambda,1}(c_k,t)-1}{r_{\lambda,1}(c_k,t)}\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\cdot\cos\left(\omega_{\lambda,1}(c_k,t)\right)\\ &\quad+\frac{1}{r_{\lambda,1}(c_k,t)}\cos\left(\epsilon_{\lambda,1}(c_k,t)\right)\cdot\frac{\sin\left(\omega_{\lambda,1}(c_k,t)\right)}{\omega_{\lambda,1}(c_k,t)}.\end{aligned}\tag{3.3.13}$$

Since $r_{\lambda,1}(c_k,t)$ is close to 1 and $\epsilon_{\lambda,1}(c_k,t)$ is close to 0 uniformly for every eigenvalue (c.f., Sections 3.9 and 3.10), the behaviour of $\boldsymbol{B}_{\lambda,1}\left(c_k,t\right)$ will be encapsulated in terms of trigonometric functions with argument $\omega_{\lambda,1}(c_k,t)$, provided some care is taken to make every singularity removable (c.f., Sections 3.9 and 3.10). As will become clear, this serves to make the behaviour and magnitude of $\boldsymbol{B}_{\lambda,1}\left(c_k,t\right)$ explicit, which, in turn, serves to reduce the number of function evaluations and volume of linear algebra in the quadrature schemes as well as to decrease the quadrature error without using derivatives of the potential.

To make this guiding principle precise, the following definition introduces the non-trigonometric parts $\boldsymbol{f}_{\lambda,1}(c_k,t)\in\mathbb{R}^{7\times1}$, $\boldsymbol{\iota}_{\lambda,1}(c_k,t)\in\mathbb{R}^{3\times1}$ and $\boldsymbol{g}_{\lambda,1}(c_k,t)\in\mathbb{R}^{7\times1}$, which appear below in Theorems 3.3.1, 3.3.2 and 3.3.3. Although important, it is technical in nature and the reader is encouraged to glance over it and return to it as needed.

Definition 3.3.2.

$$\begin{aligned}\boldsymbol{f}_{\lambda,1}(c_k,t) &:= \frac{q(t)-\frac{\int_{c_k}^{t}q(\xi)d\xi}{t-c_k}}{t-c_k}\frac{1}{r_{\lambda,1}^2(c_k,t)}\\ &\times\begin{bmatrix}1\\ \left(r_{\lambda,1}(c_k,t)-1\right)\left(r_{\lambda,1}(c_k,t)\varphi\left(i\cdot\epsilon_{\lambda,1}(c_k,t)\right)-\frac{1-\cos\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}^2(c_k,t)}\right)\\ -r_{\lambda,1}(c_k,t)\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right)-\frac{r_{\lambda,1}(c_k,t)-1}{r_{\lambda,1}(c_k,t)}\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right)\\ -2\left(r_{\lambda,1}(c_k,t)-1\right)\left(\frac{\left(r_{\lambda,1}(c_k,t)-1\right)^2}{r_{\lambda,1}(c_k,t)}\phi\left(i\cdot\epsilon_{\lambda,1}(c_k,t)\right)+\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right)\\ -\frac{2}{r_{\lambda,1}(c_k,t)}\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right)+r_{\lambda,1}(c_k,t)s_{\lambda,1}(c_k,t)\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right)\\ \frac{1}{2}r_{\lambda,1}^2(c_k,t)\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right)-\frac{r_{\lambda,1}(c_k,t)-1}{r_{\lambda,1}(c_k,t)}\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right)\\ -\frac{1}{2}r_{\lambda,1}(c_k,t)\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right)+r_{\lambda,1}(c_k,t)s_{\lambda,1}(c_k,t)\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right)\end{bmatrix},\\ \boldsymbol{f}_{\lambda,1}(c_k,c_k) &:= \frac{q'\left(c_k^+\right)}{2}\begin{bmatrix}1 & 0 & -1 & 0 & -2 & \frac{1}{2} & -\frac{1}{2}\end{bmatrix}^\top,\end{aligned}$$

$$\boldsymbol{\iota}_{\lambda,1}(c_k,t) := \frac{q(t)-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}{t-c_k}\begin{bmatrix} \varphi\left(\sqrt{\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)}\right) \\ -2\phi\left(\sqrt{\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)}\right) \\ \frac{1}{2}\phi\left(\sqrt{\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)}\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) \end{bmatrix},$$

$$\boldsymbol{\iota}_{\lambda,1}(c_k,c_k) := \frac{q'\left(c_k^+\right)}{2}\begin{bmatrix} -\frac{1}{2} & -\frac{2}{3} & 0 \end{bmatrix}^\top,$$

$$\boldsymbol{g}_{\lambda,1}(c_k,t) := \frac{1}{2}\frac{q(t)-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}{t-c_k}\frac{1}{r_{\lambda,1}^2(c_k,t)}$$

$$\times\begin{bmatrix} 1 \\ (r_{\lambda,1}(c_k,t)-1)\frac{1-\cos\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)} - r_{\lambda,1}(c_k,t)\sin\left(\epsilon_{\lambda,1}(c_k,t)\right) \\ -r_{\lambda,1}(c_k,t)\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right) - \frac{r_{\lambda,1}(c_k,t)-1}{r_{\lambda,1}(c_k,t)}\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right) \\ \cos\left(\epsilon_{\lambda,1}(c_k,t)\right) - \frac{r_{\lambda,1}(c_k,t)-1}{r_{\lambda,1}(c_k,t)}\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)} \\ -\frac{1}{r_{\lambda,1}(c_k,t)}\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right) + r_{\lambda,1}(c_k,t)s_{\lambda,1}(c_k,t)\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right) \\ r_{\lambda,1}^2(c_k,t)\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right) - \frac{r_{\lambda,1}(c_k,t)-1}{r_{\lambda,1}(c_k,t)}\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right) \\ -r_{\lambda,1}(c_k,t)\left(\cos\left(\epsilon_{\lambda,1}(c_k,t)\right) + r_{\lambda,1}(c_k,t)s_{\lambda,1}(c_k,t)\frac{\sin\left(\epsilon_{\lambda,1}(c_k,t)\right)}{\epsilon_{\lambda,1}(c_k,t)}\right) \end{bmatrix},$$

$$\boldsymbol{g}_{\lambda,1}(c_k,c_k) := \frac{q'\left(c_k^+\right)}{4}\begin{bmatrix} 1 & 0 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}^\top.$$

With Definition 3.3.2 in hand, it is now possible to write the following three theorems.

Theorem 3.3.1 (Ramos, 2015*a*). *If λ belongs to (3.3.1) then*

$$\begin{aligned}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}\left(c_k,t\right)\right) = {} & \frac{1-\cos\left(\omega_{\lambda,1}(c_k,t)\right)}{\left(\omega_{\lambda,1}(c_k,t)\right)^2}(t-c_k)\begin{bmatrix} t-c_k \\ 0 \\ 0 \end{bmatrix} \odot \begin{bmatrix} [\boldsymbol{f}_{\lambda,1}(c_k,t)]_{1,1} \\ 0 \\ 0 \end{bmatrix} \\ & + \cos\left(\omega_{\lambda,1}(c_k,t)\right)(t-c_k)\begin{bmatrix} t-c_k \\ (t-c_k)^2 \\ 1 \end{bmatrix} \odot \begin{bmatrix} [\boldsymbol{f}_{\lambda,1}(c_k,t)]_{2,1} \\ [\boldsymbol{f}_{\lambda,1}(c_k,t)]_{4,1} \\ [\boldsymbol{f}_{\lambda,1}(c_k,t)]_{6,1} \end{bmatrix} \\ & + \frac{\sin\left(\omega_{\lambda,1}(c_k,t)\right)}{\omega_{\lambda,1}(c_k,t)}(t-c_k)\begin{bmatrix} t-c_k \\ 0 \\ 1 \end{bmatrix} \odot \begin{bmatrix} [\boldsymbol{f}_{\lambda,1}(c_k,t)]_{3,1} \\ 0 \\ [\boldsymbol{f}_{\lambda,1}(c_k,t)]_{7,1} \end{bmatrix} \\ & + \phi\left(i\cdot\omega_{\lambda,1}(c_k,t)\right)(t-c_k)\begin{bmatrix} 0 \\ (t-c_k)^2 \\ 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ [\boldsymbol{f}_{\lambda,1}(c_k,t)]_{5,1} \\ 0 \end{bmatrix}\end{aligned}$$

where $\omega_{\lambda,1}(c_k,t)$ and $\boldsymbol{f}_{\lambda,1}(c_k,t)$ are as in Definitions 3.3.1–3.3.2. Furthermore, the derivatives $\boldsymbol{f}^{(j)}_{\lambda,1}(c_k,t)$ can be bounded independently of λ.

Proof. See Section 3.9. □

Theorem 3.3.2 (Ramos, 2015*a*). *It is true that*

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}\left(c_k,t\right)\right) = (t-c_k)\begin{bmatrix} t-c_k \\ (t-c_k)^2 \\ 1 \end{bmatrix} \odot \boldsymbol{\iota}_{\lambda,1}(c_k,t)$$

where $\boldsymbol{\iota}_{\lambda,1}(c_k,t)$ is as in Definition 3.3.2. Furthermore, if λ belongs to (3.3.2) then the derivatives $\boldsymbol{\iota}^{(j)}_{\lambda,1}(c_k,t)$ can be bounded independently of λ.

Proof. Follows immediately from Theorem 3.2.2. □

Theorem 3.3.3 (Ramos, 2015*a*). *If λ belongs to* (3.3.3) *then*

$$\begin{aligned}
\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}\left(c_k,t\right)\right) = {} & \frac{1-\cos\left(\omega_{\lambda,1}(c_k,t)\right)}{\omega_{\lambda,1}(c_k,t)}(t-c_k)\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}}\\0\\0\end{bmatrix}\odot\begin{bmatrix}[\boldsymbol{g}_{\lambda,1}(c_k,t)]_{1,1}\\0\\0\end{bmatrix}\\
& +\cos\left(\omega_{\lambda,1}(c_k,t)\right)(t-c_k)\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}}\\\frac{1}{\lambda-q(c_k)}\\1\end{bmatrix}\odot\begin{bmatrix}[\boldsymbol{g}_{\lambda,1}(c_k,t)]_{2,1}\\[\boldsymbol{g}_{\lambda,1}(c_k,t)]_{4,1}\\[\boldsymbol{g}_{\lambda,1}(c_k,t)]_{6,1}\end{bmatrix}\\
& +\frac{\sin\left(\omega_{\lambda,1}(c_k,t)\right)}{\omega_{\lambda,1}(c_k,t)}(t-c_k)\begin{bmatrix}0\\\frac{1}{\lambda-q(c_k)}\\1\end{bmatrix}\odot\begin{bmatrix}0\\[\boldsymbol{g}_{\lambda,1}(c_k,t)]_{5,1}\\[\boldsymbol{g}_{\lambda,1}(c_k,t)]_{7,1}\end{bmatrix}\\
& +\sin\left(\omega_{\lambda,1}(c_k,t)\right)(t-c_k)\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}}\\0\\0\end{bmatrix}\odot\begin{bmatrix}[\boldsymbol{g}_{\lambda,1}(c_k,t)]_{3,1}\\0\\0\end{bmatrix}
\end{aligned}$$

where $\omega_{\lambda,1}(c_k,t)$ and $\boldsymbol{g}_{\lambda,1}(c_k,t)$ are as in Definitions 3.3.1–3.3.2. Furthermore, the derivatives $\boldsymbol{g}^{(j)}_{\lambda,1}(c_k,t)$ can be bounded independently of λ.

Proof. See Section 3.10. □

3.3.2 Exploiting the magnitude to reduce the number of function evaluations and volume of linear algebra

The following definition, corollaries and theorems serve to illustrate that it is possible to use the representations in Theorems 3.3.1, 3.3.2 and 3.3.3 to develop a quadrature which exploits the magnitude of each integrand in order to reduce the number of function evaluations and volume of linear algebra. In particular, Definition 3.3.3 below introduces $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k+h_kt)$ as a means to decompose the fine and coarse parts of $\boldsymbol{B}_{\lambda,1}(c_k, c_k+h_kt)$. This fine and coarse decomposition is made precise with Corollaries 3.3.1, 3.3.2 and 3.3.3 below which show that $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k+h_kt)$ is $\mathcal{O}(1)$ uniformly over the entire eigenvalue range. It is then in Theorems 3.3.4, 3.3.5, 3.3.6, 3.3.7 and 3.3.8 that this fine and coarse decomposition is shown to bear fruit in the form of fewer function evaluations and linear algebra.

Definition 3.3.3. *Let* $\boldsymbol{B}^{fine}_{\lambda,1}(c_k, c_k + h_k t)$ *be the unique element in* $\mathfrak{sl}(2,\mathbb{R})$ *such that*

$$\begin{aligned}
&\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k, c_k + h_k t)\right) \\
&=: \boldsymbol{\pi}\left(\boldsymbol{B}^{fine}_{\lambda,1}(c_k, c_k + h_k t)\right) \odot \begin{cases} h_k \begin{bmatrix} h_k & h_k^2 & 1 \end{bmatrix}^\top, & |\lambda - q_{max}| \leq h_{max}^{-2}, \\ h_k \begin{bmatrix} \frac{1}{\sqrt{\lambda - q(c_k)}} & \frac{1}{\lambda - q(c_k)} & 1 \end{bmatrix}^\top, & \lambda - q_{max} \geq h_{max}^{-2}. \end{cases}
\end{aligned}$$

Corollary 3.3.1 (Ramos, 2015*a*)**.** *If* λ *lies in* (3.3.1) *then either* $\omega_{\lambda,1}(c_k, c_{k+1}) \in \left[0, 2i\right]$ *or* $\omega_{\lambda,1}(c_k, c_{k+1}) \in \left[0, 2\sqrt{2}\right]$ *and*

$$\begin{aligned}
\boldsymbol{\pi}\left(\boldsymbol{B}^{fine}_{\lambda,1}\left(c_k, c_k + h_k t\right)\right) = &\frac{1 - \cos\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right)}{\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right)^2} \begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix} \odot t \begin{bmatrix} [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{1,1} \\ 0 \\ 0 \end{bmatrix} \\
&+ \cos\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right) \begin{bmatrix} t \\ t^2 \\ 1 \end{bmatrix} \odot t \begin{bmatrix} [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{2,1} \\ [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{4,1} \\ [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{6,1} \end{bmatrix} \\
&+ \frac{\sin\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right)}{\omega_{\lambda,1}(c_k, c_{k+1})t} \begin{bmatrix} t \\ 0 \\ 1 \end{bmatrix} \odot t \begin{bmatrix} [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{3,1} \\ 0 \\ [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{7,1} \end{bmatrix} \\
&+ \phi\left(i \cdot \omega_{\lambda,1}(c_k, c_{k+1})t\right) \begin{bmatrix} 0 \\ t^2 \\ 0 \end{bmatrix} \odot t \begin{bmatrix} 0 \\ [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{5,1} \\ 0 \end{bmatrix}.
\end{aligned} \tag{3.3.14}$$

Proof. Follows immediately from Theorem 3.3.1. □

Corollary 3.3.2 (Ramos, 2015*a*)**.** *If* λ *lies in* (3.3.2) *then*

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}^{fine}\left(c_k, c_k + h_k t\right)\right) = \begin{bmatrix} t \\ t^2 \\ 1 \end{bmatrix} \odot t \cdot \begin{bmatrix} [\boldsymbol{\iota}_{\lambda,1}(c_k, c_k + h_k t)]_{1,1} \\ [\boldsymbol{\iota}_{\lambda,1}(c_k, c_k + h_k t)]_{2,1} \\ [\boldsymbol{\iota}_{\lambda,1}(c_k, c_k + h_k t)]_{3,1} \end{bmatrix}. \tag{3.3.15}$$

Proof. Follows immediately from Theorem 3.3.2. □

Corollary 3.3.3 (Ramos, 2015*a*)**.** *If* λ *lies in* (3.3.3) *then* $\omega_{\lambda,1}(c_k, c_{k+1}) \in [1, +\infty)$ *and*

$$\begin{aligned} \boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}^{fine}\left(c_k, c_k + h_k t\right)\right) = {} & \frac{1 - \cos\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right)}{\omega_{\lambda,1}(c_k, c_{k+1})t} t \begin{bmatrix} [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{1,1} \\ 0 \\ 0 \end{bmatrix} \\ & + \cos\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right) t \begin{bmatrix} [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{2,1} \\ [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{4,1} \\ [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{6,1} \end{bmatrix} \\ & + \frac{\sin\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right)}{\omega_{\lambda,1}(c_k, c_{k+1})t} t \begin{bmatrix} 0 \\ [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{5,1} \\ [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{7,1} \end{bmatrix} \\ & + \sin\left(\omega_{\lambda,1}(c_k, c_{k+1})t\right) t \begin{bmatrix} [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{3,1} \\ 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{3.3.16}$$

Proof. Follows immediately from Theorem 3.3.3. □

Theorem 3.3.4 (Ramos, 2015*a*). *If λ belongs to (3.3.1) (3.3.2) or (3.3.3), then*

$$\begin{aligned}\boldsymbol{\pi}\left(\int_{c_k}^{c_{k+1}} \boldsymbol{B}_{\lambda,1}(c_k,t)dt\right) &= h_k\int_0^1 \boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt)\right)dt\\ &= \boldsymbol{\pi}\left(\int_0^1 \boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt)dt\right)\\ &\quad\odot\begin{cases} h_k^2\begin{bmatrix}h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ h_k^2\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2}.\end{cases}\end{aligned}$$

Proof. Follows by straightforward computation. □

Theorem 3.3.5 (Ramos, 2015*a*). *If λ belongs to (3.3.1) (3.3.2) or (3.3.3), then*

$$\begin{aligned}&\boldsymbol{\pi}\left(-\frac{1}{2}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\left[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)\right]dt_2dt_1\right)\\ &= -\frac{1}{2}h_k^2\int_0^1\int_0^{t_1}\boldsymbol{\pi}\left(\left[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_2),\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_1)\right]\right)dt_2dt_1\\ &= \boldsymbol{\pi}\left(-\frac{1}{2}\int_0^1\int_0^{t_1}\left[\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_2),\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_1)\right]dt_2dt_1\right)\\ &\quad\odot\begin{cases} h_k^5\begin{bmatrix}h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ h_k^4\left(\lambda-q\left(c_k\right)\right)^{-\frac{1}{2}}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2}.\end{cases}\end{aligned}$$

Proof. Follows by straightforward computation. □

Theorem 3.3.6 (Ramos, 2015*a*). *If λ belongs to (3.3.1) (3.3.2) or (3.3.3), then*

$$\begin{aligned}&\boldsymbol{\pi}\Bigg(\frac{1}{3}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\Big[\boldsymbol{B}_{\lambda,1}(c_k,t_3),\\ &\qquad\Big[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\\ &\qquad\quad\boldsymbol{B}_{\lambda,1}(c_k,t_1)\Big]\Big]dt_3dt_2dt_1\Bigg)\\ &= \frac{1}{3}h_k^3\int_0^1\int_0^{t_1}\int_0^{t_1}\boldsymbol{\pi}\Big(\Big[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_3),\\ &\qquad\Big[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_2),\\ &\qquad\quad\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_1)\Big]\Big]\Big)dt_3dt_2dt_1\end{aligned}$$

$$= \boldsymbol{\pi}\Bigg(\frac{1}{3}\int_0^1\int_0^{t_1}\int_0^{t_1}\Big[\boldsymbol{B}^{fine}_{\lambda,1}(c_k, c_k+h_k t_3),\Big[\boldsymbol{B}^{fine}_{\lambda,1}(c_k, c_k+h_k t_2),\boldsymbol{B}^{fine}_{\lambda,1}(c_k, c_k+h_k t_1)\Big]\Big]dt_3dt_2dt_1\Bigg)$$
$$\odot\begin{cases}h_k^8\begin{bmatrix}h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\le h_{max}^{-2},\\ h_k^6\left(\lambda-q\left(c_k\right)\right)^{-1}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\ge h_{max}^{-2}.\end{cases}$$

Proof. Follows by straightforward computation. □

Theorem 3.3.7 (Ramos, 2015*a*). *If λ belongs to (3.3.1) (3.3.2) or (3.3.3), then*

$$\boldsymbol{\pi}\Bigg(-\frac{1}{8}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\Big[\boldsymbol{B}_{\lambda,1}(c_k,t_4),\Big[\boldsymbol{B}_{\lambda,1}(c_k,t_3),\Big[\boldsymbol{B}_{\lambda,1}(c_k,t_2),\boldsymbol{B}_{\lambda,1}(c_k,t_1)\Big]\Big]\Big]dt_4dt_3dt_2dt_1\Bigg)$$
$$= -\frac{1}{8}h_k^4\int_0^1\int_0^{t_1}\int_0^{t_1}\int_0^{t_1}\boldsymbol{\pi}\Big(\Big[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_4),\Big[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_3),\Big[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_2),\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_1)\Big]\Big]\Big]\Big)dt_4dt_3dt_2dt_1$$
$$= \boldsymbol{\pi}\Bigg(-\frac{1}{8}\int_0^1\int_0^{t_1}\int_0^{t_1}\int_0^{t_1}\Big[\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_4),\Big[\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_3),\Big[\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_2),\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_1)\Big]\Big]\Big]dt_4dt_3dt_2dt_1\Bigg)$$
$$\odot\begin{cases}h_k^{11}\begin{bmatrix}h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\le h_{max}^{-2},\\ h_k^8\left(\lambda-q\left(c_k\right)\right)^{-\frac{3}{2}}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\ge h_{max}^{-2}.\end{cases}$$

Proof. Follows by straightforward computation. □

Theorem 3.3.8 (Ramos, 2015*a*). *If λ belongs to (3.3.1) (3.3.2) or (3.3.3), then*

$$\begin{aligned}
&\boldsymbol{\pi}\Bigg(-\frac{1}{8}\int_{c_k}^{c_{k+1}}\int_{c_k}^{t_1}\int_{c_k}^{t_1}\int_{c_k}^{t_2}\Big[\Big[\boldsymbol{B}_{\lambda,1}(c_k,t_4),\\
&\qquad\qquad\boldsymbol{B}_{\lambda,1}(c_k,t_2)\Big],\\
&\qquad\Big[\boldsymbol{B}_{\lambda,1}(c_k,t_3),\\
&\qquad\boldsymbol{B}_{\lambda,1}(c_k,t_1)\Big]\Big]dt_4dt_3dt_2dt_1\Bigg)\\
&=-\frac{1}{8}h_k^4\int_0^1\int_0^{t_1}\int_0^{t_1}\int_0^{t_2}\boldsymbol{\pi}\Big(\Big[\Big[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_4),\\
&\qquad\qquad\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_2)\Big],\\
&\qquad\Big[\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_3),\\
&\qquad\boldsymbol{B}_{\lambda,1}(c_k,c_k+h_kt_1)\Big]\Big]\Big)dt_4dt_3dt_2dt_1\\
&=\boldsymbol{\pi}\Bigg(-\frac{1}{8}\int_0^1\int_0^{t_1}\int_0^{t_1}\int_0^{t_2}\Big[\Big[\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_4),\\
&\qquad\qquad\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_2)\Big],\\
&\qquad\Big[\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_3),\\
&\qquad\boldsymbol{B}^{fine}_{\lambda,1}(c_k,c_k+h_kt_1)\Big]\Big]dt_4dt_3dt_2dt_1\Bigg)\\
&\odot\begin{cases}h_k^{11}\begin{bmatrix}h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ h_k^8\left(\lambda-q\left(c_k\right)\right)^{-\frac{3}{2}}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2}.\end{cases}
\end{aligned}$$

Proof. Follows by straightforward computation. □

Definition 3.3.3 together with Corollaries 3.3.1, 3.3.2 and 3.3.3 as well as Theorems 3.3.4, 3.3.5, 3.3.6, 3.3.7 and 3.3.8 serve to highlight the synergy between the Lie bracket and the representations in Theorems 3.3.1, 3.3.2 and 3.3.3: they act together to decrease the magnitude of each multivariate integral, making it smaller than expected!

It is of note that this would not be possible with the representations in Theorems 3.2.1, 3.2.2 and 3.2.3 because they represent $\boldsymbol{B}_{\lambda,1}(c_k,t)$ in terms of larger quantities.

As a result of this interaction, the quadrature of (3.1.1)–(3.1.5) should be replaced with that of

$$\int_0^1\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k,c_k+h_kt)dt, \tag{3.3.17}$$

$$-\frac{1}{2}\int_0^1\int_0^{t_1}\left[\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k,c_k+h_kt_2),\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k,c_k+h_kt_1)\right]dt_2dt_1, \tag{3.3.18}$$

$$\frac{1}{3}\int_0^1\int_0^{t_1}\int_0^{t_1}\left[\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_3),\left[\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_2),\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_1)\right]\right]dt_3dt_2dt_1, \tag{3.3.19}$$

$$-\frac{1}{8}\int_0^1\int_0^{t_1}\int_0^{t_1}\int_0^{t_1}\Big[\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_4),\Big[\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_3),\Big[\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_2),\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_1)\Big]\Big]\Big]dt_4dt_3dt_2dt_1, \tag{3.3.20}$$

$$-\frac{1}{8}\int_0^1\int_0^{t_1}\int_0^{t_1}\int_0^{t_2}\Big[\Big[\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_4),\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_2)\Big],\Big[\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_3),\boldsymbol{B}_{\lambda,1}^{\text{fine}}(c_k,c_k+h_kt_1)\Big]\Big]dt_4dt_3dt_2dt_1, \tag{3.3.21}$$

since this results in fewer function evaluations and volume of linear algebra, as described in the next four subsubsections.

3.3.2.1 Global order 4

According to Theorem 3.3.4, the quadrature of (3.1.1) up to local order 5, is equivalent to the quadrature of (3.3.17) up to local order 3.

3.3.2.2 Global order 7

As a consequence of Theorem 3.3.4, it is immediate that the quadrature of (3.1.1) up to local order 8, is equivalent to the quadrature of (3.3.17) up to local order 6. Following Theorem 3.3.5 and the fact that

$$\begin{aligned}
\lambda\in[q_{\max}+h_{\max}^{-2},q_{\max}+h_{\max}^{-4})&\Rightarrow h_k^4\left(\lambda-q\left(c_k\right)\right)^{-\frac{1}{2}}\leq h_{\max}^5,\\
\lambda\in[q_{\max}+h_{\max}^{-4},q_{\max}+h_{\max}^{-6})&\Rightarrow h_k^4\left(\lambda-q\left(c_k\right)\right)^{-\frac{1}{2}}\leq h_{\max}^6,\\
\lambda\in[q_{\max}+h_{\max}^{-6},q_{\max}+h_{\max}^{-8})&\Rightarrow h_k^4\left(\lambda-q\left(c_k\right)\right)^{-\frac{1}{2}}\leq h_{\max}^7,\\
\lambda\in[q_{\max}+h_{\max}^{-8},+\infty)&\Rightarrow h_k^4\left(\lambda-q\left(c_k\right)\right)^{-\frac{1}{2}}\leq h_{\max}^8,
\end{aligned}$$

it is clear that the quadrature of (3.1.2) up to local order 8, is equivalent to the quadrature of (3.3.18) up to local order 3, 2, 1, 0, for

$$\begin{aligned}
&\lambda \in [q_{\max} - h_{\max}^{-2}, q_{\max} + h_{\max}^{-4}),\\
&\lambda \in [q_{\max} + h_{\max}^{-4}, q_{\max} + h_{\max}^{-6}),\\
&\lambda \in [q_{\max} + h_{\max}^{-6}, q_{\max} + h_{\max}^{-8}),\\
&\lambda \in [q_{\max} + h_{\max}^{-8}, +\infty),
\end{aligned}$$

respectively. This quantifies the "at most" feature described in Subsubsection 1.1.5.3 for this case.

3.3.2.3 Global order 10

Similarly to the case of global order 7, *i*) Theorem 3.3.4 guarantees that the quadrature of (3.1.1) up to local order 11 is equivalent to the quadrature of (3.3.17) up to local order 9, *ii*) Theorem 3.3.5 ensures that the quadrature of (3.1.2) up to local order 11, is equivalent to the quadrature of (3.3.18) up to at most local order 6, and, *iii*) Theorem 3.3.6 establishes that the quadrature of (3.1.3) up to local order 11, is equivalent to the quadrature of (3.3.19) up to at most local order 3.

3.3.2.4 Global order 13

Analogously to the case of global order 10, *i*) Theorem 3.3.4 guarantees that the quadrature of (3.1.1) up to local order 14 is equivalent to the quadrature of (3.3.17) up to local order 12, *ii*) Theorem 3.3.5 ensures that the quadrature of (3.1.2) up to local order 14, is equivalent to the quadrature of (3.3.18) up to at most local order 9, *iii*) Theorem 3.3.6 establishes that the quadrature of (3.1.3) up to local order 14, is equivalent to the quadrature of (3.3.19) up to at most local order 6, and, *iv*) Theorems 3.3.7–3.3.8 establish that the quadrature of (3.1.4)–(3.1.5) up to local order 14, is equivalent to the quadrature of (3.3.20)–(3.3.21) up to at most local order 3.

3.3.3 Exploiting the behaviour to decrease the quadrature error without using derivatives of the potential

Very much along the same lines as in the theoretical analysis in (Levin, 1996; Iserles and Nørsett, 2005; Iserles and Nørsett, 2006) for highly oscillatory Fourier-type integrands, the mildly to highly oscillatory unconventional-type behaviour for $\lambda \in \left[q_{\max} + h_{\max}^{-2}, +\infty\right)$ made explicit in Corollary 3.3.3 also suggests a polynomial interpolation which includes the endpoints. This is so because it makes the difference between the function and the interpolation polynomial equal to zero at the endpoints, which, in many cases, can be shown to result in a decrease in quadrature error. This would lead to the evaluation of

$\boldsymbol{g}_{\lambda,1}(c_k, c_k)$, which depends on $q'\left(c_k^+\right)$ (c.f., Definition 3.3.2): something that would be best to avoid since the derivative of the potential might not be available in closed-form. Fortunately, since in Corollary 3.3.3 there is a 't' term in front of every '$\left[\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)\right]_{j,1}$' term, this is automatically achieved at the left boundary point. Hence, there is no need to interpolate at the left boundary point. As for $t \in (0, 1]$, $\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)$ does not depend on the derivative of the potential and should be interpolated at the right boundary point.

This makes it possible to decrease the quadrature error without using derivatives of the potential, and would not be possible with the representations in Theorems 3.2.1 and 3.2.3 because they represent $\boldsymbol{B}_{\lambda,1}(c_k, c_k)$ as a sum where each term is a product of a non-zero vector times $q'\left(c_k^+\right)$.

3.3.4 Optimal interpolation

In view of Corollaries 3.3.1, 3.3.2 and 3.3.3, the selection of a quadrature which respects the behaviour of each integrand is possible by a polynomial interpolation of $\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)$, $\boldsymbol{\iota}_{\lambda,1}(c_k, c_k + h_k t)$ and $\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)$ in $t \in [0, 1]$ and the exact integration of the result: a Filon-type quadrature. Thus, multivariate quadrature over polytopes becomes univariate polynomial interpolation over intervals.

The results in this subsection focus on $t \mapsto \boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)$, but they also hold verbatim for $t \mapsto \boldsymbol{\iota}_{\lambda,1}(c_k, c_k + h_k t)$ and $t \mapsto \boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)$.

3.3.4.1 Smallest number of interpolation points to be consistent with local order

Let $\tau_1, \tau_2, \cdots, \tau_{j-1}, \tau_j$ be j interpolation points such that

$$0 \leq \tau_1 < \tau_2 < \cdots < \tau_{j-1} < \tau_j \leq 1$$

and let $t \mapsto \boldsymbol{p}_{\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k \cdot), j-1}(t)$ be the unique (at most) $j-1$ degree interpolation polynomial such that, for every $l \in \{1, 2, \cdots, j-1, j\}$,

$$\boldsymbol{p}_{\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k \cdot), j-1}(\tau_l) = \boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k \tau_l).$$

Then (Olver, Lozier, Boisvert and Clark, 2010, Subsection 3.3(i)), for every $t \in [0, 1]$, there exists $\xi \in [0, 1]$ such that

$$\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t) - \boldsymbol{p}_{\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k \cdot), j-1}(t) = \frac{h_k^j \boldsymbol{f}_{\lambda,1}^{(j)}(c_k, c_k + h_k \xi)}{j!} \prod_{l=1}^{j} (t - \tau_l)$$

which, in turn, yields the pointwise error estimate for polynomial interpolation: for every $t \in [0,1]$,

$$\begin{aligned}&\left| t\left[\boldsymbol{f}_{\lambda,1}(c_k, c_k+h_k t) - \boldsymbol{p}_{\boldsymbol{f}_{\lambda,1}(c_k,c_k+h_k\cdot),j-1}(t) \right]_{i,1} \right| \\ &\leq \frac{h_k^j}{j!} \max_{\xi\in[c_k,c_{k+1}]} \left\{ \left| \left[\boldsymbol{f}_{\lambda,1}^{(j)}(c_k,\xi) \right]_{i,1} \right| \right\} \max_{\xi\in[0,1]} \left\{ \left| \xi \prod_{l=1}^{j} (\xi-\tau_l) \right| \right\}. \end{aligned} \tag{3.3.22}$$

Together with the discussion at the end of Subsection 3.3.2, (3.3.22) dictates that *i*) global order 4 requires $j=3$ for (3.3.17), *ii*) global order 7 requires $j=6$ for (3.3.17) and $j=3$ for (3.3.18), *iii*) global order 10 requires $j=9$ for (3.3.17), $j=6$ for (3.3.18) and $j=3$ for (3.3.19), and, *iv*) global order 13 requires $j=12$ for (3.3.17), $j=9$ for (3.3.18), $j=6$ for (3.3.19) and $j=3$ for (3.3.20)–(3.3.21).

Paradoxically, fewer interpolation points are needed for higher dimensional integrals than for lower dimensional integrals, which represents a huge saving in function evaluations and linear algebra!

3.3.4.2 Interpolation points that decrease the quadrature error without using derivatives of the potential

As discussed in Subsection 3.3.3, in order to decrease the quadrature error without using derivatives of the potential, interpolate $t \mapsto \boldsymbol{f}_{\lambda,1}(c_k, c_k+h_k t)$, $t \mapsto \boldsymbol{\iota}_{\lambda,1}(c_k, c_k+h_k t)$ and $t \mapsto \boldsymbol{g}_{\lambda,1}(c_k, c_k+h_k t)$ at $t=1$, but not at $t=0$, i.e., choose $\tau_1 :\neq 0$ and $\tau_j := 1$, since, in this case, $t \mapsto t\boldsymbol{p}_{\boldsymbol{f}_{\lambda,1}(c_k,c_k+h_k\cdot),j-1}(t)$ is the unique (at most) j degree interpolation polynomial that interpolates $t \mapsto t\boldsymbol{f}_{\lambda,1}(c_k, c_k+h_k t)$ at the $j+1$ points $\{0, \tau_1, \cdots, \tau_{j-1}, 1\}$.

3.3.4.3 Data

The polynomial interpolation in this section requires the following data

$$\bigcup_{k=0}^{m-1} \left(\{q(c_k)\} \cup \left\{ q(c_k+h_k t), \int_{c_k}^{c_k+h_k t} q(\xi)d\xi : t \in \mathcal{S} \right\} \cup \left\{ \int_{c_k}^{c_{k+1}} q(\xi)d\xi \right\} \right) \cup \{q(b)\}$$

where

$$\mathcal{S} := \begin{cases} \{\tau_1, \tau_2\}, & \text{for global order 4,} \\ \{\tau_1, \tau_2, \tau_3, \tau_4, \tau_5\}, & \text{for global order 7,} \\ \{\tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6, \tau_7, \tau_8\}, & \text{for global order 10,} \\ \{\tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6, \tau_7, \tau_8, \tau_9, \tau_{10}, \tau_{11}\}, & \text{for global order 13.} \end{cases}$$

If the antiderivative of the potential is not available in closed-form, then it is possible to approximate, up to local order, the antiderivative data

$$\left\{\int_{c_k}^{c_k+h_k t} q(\xi)d\xi : t \in \mathcal{S}\right\} \cup \left\{\int_{c_k}^{c_{k+1}} q(\xi)d\xi\right\}$$

by the polynomial interpolation of $q(\xi)$ in $\xi \in [c_k, c_{k+1}]$ with the potential data

$$\{q(c_k)\} \cup \{q(c_k + h_k t) : t \in \mathcal{S}\} \cup \{q(c_{k+1})\}$$

and the exact integration of the result.

3.4 Error estimates

The current section quantifies the total error in the Fer streamers approach to Sturm–Liouville problems. To this end, Definition 3.4.1 and Theorem 3.4.1 below make explicit the quadrature error in Section 3.3, and Definition 3.4.2, Theorem 3.4.2 and Corollary 3.4.1 below clarify the manner in which the quadrature error, which lives in the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$, affects the various quantities in the Lie group $\mathrm{SL}(2,\mathbb{R})$. Finally, Definition 3.4.3 and Theorem 3.4.3 below quantify the total error in the Fer streamers approach, in terms of the truncation estimates in Corollary 2.1.1 and the discretization estimates in the present chapter.

Definition 3.4.1. *For $n = 1, \log(3)/\log(2), 2, \log(5)/\log(2)$, i.e., global order $4, 7, 10, 13$, let*

$$\tilde{\boldsymbol{D}}_{\lambda,1,n}(c_k, c_{k+1})$$

denote the approximation in $\mathfrak{sl}(2,\mathbb{R})$ of the univariate integral (3.1.1) in $\boldsymbol{D}_{\lambda,1}(c_k, c_{k+1})$ with the optimal quadrature in Section 3.3, using $3, 6, 9, 12$ interpolation points, respectively. Let also

$$\boldsymbol{E}_{\lambda,1,n}(c_k, c_{k+1}) := \boldsymbol{D}_{\lambda,1}(c_k, c_{k+1}) - \tilde{\boldsymbol{D}}_{\lambda,1,n}(c_k, c_{k+1})$$

denote the error in that approximation. For global order 7, let

$$\tilde{\boldsymbol{D}}_{\lambda,2,\log(3)/\log(2)}(c_k, c_{k+1})$$

denote the approximation in $\mathfrak{sl}(2,\mathbb{R})$ of the bivariate integral (3.1.2) in $\boldsymbol{D}_{\lambda,2}(c_k, c_{k+1})$ with the optimal quadrature in Section 3.3 with 3 interpolation points. For global order 10, let

$$\tilde{\boldsymbol{D}}_{\lambda,2,2}(c_k, c_{k+1})$$

denote the approximation in $\mathfrak{sl}(2,\mathbb{R})$ of the bivariate (3.1.2) and trivariate (3.1.3) integrals in $\boldsymbol{D}_{\lambda,2}(c_k, c_{k+1})$ with the optimal quadrature in Section 3.3 with 6 and 3 interpolation

points, respectively. For global order 13*, let*

$$\tilde{\boldsymbol{D}}_{\lambda,2,\log(5)/\log(2)}(c_k, c_{k+1})$$

denote the approximation in $\mathfrak{sl}(2,\mathbb{R})$ *of the bivariate* (3.1.2)*, trivariate* (3.1.3) *and quadrivariate* (3.1.4) *integrals in* $\boldsymbol{D}_{\lambda,2}(c_k, c_{k+1})$ *with the optimal quadrature in Section 3.3 with* 9*,* 6 *and* 3 *interpolation points, respectively. Also, for* $n \in \{\log(3)/\log(2), 2, \log(5)/\log(2)\}$*, let*

$$\boldsymbol{E}_{\lambda,2,n}(c_k, c_{k+1}) := \boldsymbol{D}_{\lambda,2}(c_k, c_{k+1}) - \tilde{\boldsymbol{D}}_{\lambda,2,n}(c_k, c_{k+1})$$

denote the error in that approximation. For global order 13*, let*

$$\tilde{\boldsymbol{D}}_{\lambda,3,\log(5)/\log(2)}(c_k, c_{k+1})$$

denote the approximation in $\mathfrak{sl}(2,\mathbb{R})$ *of the quadrivariate* (3.1.5) *integral in* $\boldsymbol{D}_{\lambda,3}(c_k, c_{k+1})$ *with the optimal quadrature in Section 3.3 with* 3 *interpolation points. Finally, for* $n = \log(5)/\log(2)$*, let*

$$\boldsymbol{E}_{\lambda,3,\log(5)/\log(2)}(c_k, c_{k+1}) := \boldsymbol{D}_{\lambda,3}(c_k, c_{k+1}) - \tilde{\boldsymbol{D}}_{\lambda,3,\log(5)/\log(2)}(c_k, c_{k+1})$$

denote the error in that approximation.

Theorem 3.4.1 (Ramos, 2015*a*)**.** *If Assumption 1.1.1 holds true: if* $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$*, then*

$$\boldsymbol{\pi}\left(\boldsymbol{E}_{\lambda,1,n}(c_k, c_{k+1})\right) = h_{\max}^{3\times 2^n - 1} \begin{cases} \begin{bmatrix} \mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top, & w.r.t\ (1.1.6), \\ \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} & \mathcal{O}\left(1\right) \end{bmatrix}^\top, & w.r.t\ (1.1.7), \end{cases}$$

whereas, if $n \in \{\log(3)/\log(2), 2, \log(5)/\log(2)\}$*, then*

$$\boldsymbol{\pi}\left(\boldsymbol{E}_{\lambda,2,n}(c_k, c_{k+1})\right) = \mathcal{O}\left(\boldsymbol{\pi}\left(\boldsymbol{E}_{\lambda,1,n}(c_k, c_{k+1})\right)\right),$$

and, if $n = \log(5)/\log(2)$*, then*

$$\boldsymbol{\pi}\left(\boldsymbol{E}_{\lambda,3,n}(c_k, c_{k+1})\right) = \mathcal{O}\left(\boldsymbol{\pi}\left(\boldsymbol{E}_{\lambda,1,n}(c_k, c_{k+1})\right)\right).$$

Proof. Follows from the discussion in Section 3.1, from Theorems 3.3.4, 3.3.5, 3.3.6, 3.3.7 and 3.3.8, and from the discussion in Subsection 3.3.4. □

Definition 3.4.2. *Let $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$, and define the discretized flow, the discretized solution, the discretization local error, and the discretization global error by*

$$\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1}) := \begin{cases} e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{D}}_{\lambda,1,1}(c_k,c_{k+1})} \Leftarrow n = 1, \\ e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{D}}_{\lambda,1,n}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{D}}_{\lambda,2,n}(c_k,c_{k+1})} \Leftarrow n \in \left\{ \dfrac{\log(3)}{\log(2)}, 2 \right\}, \\ e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{D}}_{\lambda,1,n}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{D}}_{\lambda,2,n}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{D}}_{\lambda,3,n}(c_k,c_{k+1})} \Leftarrow n = \dfrac{\log(5)}{\log(2)}, \end{cases}$$

$$\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1}) := \tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1}) \cdots \tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_1, c_2) \tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(a, c_1),$$

$$\boldsymbol{L}^{disc.}_{\lambda,n}(c_k, c_{k+1}) := \log\left(\tilde{\boldsymbol{F}}_{\lambda,\lceil n \rceil}(c_k, c_{k+1}) \tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k, c_{k+1}) \right),$$

$$\boldsymbol{G}^{disc.}_{\lambda,n}(c_{k+1}) := \log\left(\tilde{\boldsymbol{Y}}_{\lambda,\lceil n \rceil}(c_{k+1}) \tilde{\tilde{\boldsymbol{Y}}}^{-1}_{\lambda,n}(c_{k+1}) \right),$$

respectively.

Theorem 3.4.2 (Ramos, 2015*a*)**.** *If $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$ and Assumption 1.1.1 holds true, then*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{disc.}_{\lambda,n}(c_k, c_{k+1}) \right) = h_{\max}^{3\times 2^n - 1} \begin{cases} \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}(1) \end{bmatrix}^\top, \ w.r.t \ (1.1.6), \\ \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} & \mathcal{O}(1) \end{bmatrix}^\top, \ w.r.t \ (1.1.7), \end{cases}$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{disc.}_{\lambda,n}(c_{k+1}) \right) = h_{\max}^{3\times 2^n - 2} \begin{cases} \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}(1) \end{bmatrix}^\top, \ w.r.t \ (1.1.6), \\ \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} & \mathcal{O}(1) \end{bmatrix}^\top, \ w.r.t \ (1.1.7). \end{cases}$$

Proof. See Section 3.11. □

Corollary 3.4.1 (Ramos, 2015*a*)**.** *If $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$ and Assumption 1.1.1 holds true, then, in the two uniform regimes (1.1.6) and (1.1.7),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{disc.}_{\lambda,n}(c_k, c_{k+1}) \right) = h_{\max}^{3\times 2^n - 1} \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}(1) \end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{disc.}_{\lambda,n}(c_{k+1}) \right) = h_{\max}^{3\times 2^n - 2} \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}(1) \end{bmatrix}^\top.$$

Definition 3.4.3. *Let* $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$*, and define the*

$$\begin{aligned} &\textit{total local error:} && \boldsymbol{L}^{total}_{\lambda,n}(c_k, c_{k+1}) := \log\left(\boldsymbol{F}_\lambda(c_k, c_{k+1})\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k, c_{k+1})\right), \\ &\textit{total global error:} && \boldsymbol{G}^{total}_{\lambda,n}(c_{k+1}) := \log\left(\boldsymbol{Y}_\lambda(c_{k+1})\tilde{\tilde{\boldsymbol{Y}}}^{-1}_{\lambda,n}(c_{k+1})\right). \end{aligned}$$

Theorem 3.4.3 (Ramos, 2015*a*)**.** *If* $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$ *and Assumption 1.1.1 holds true, then, in the two uniform regimes (1.1.6) and (1.1.7),*

$$\begin{aligned} \boldsymbol{L}^{total}_{\lambda,n}(c_k, c_{k+1}) &= \boldsymbol{L}^{trun.}_{\lambda,\lceil n\rceil}(c_k, c_{k+1}) + \boldsymbol{L}^{disc.}_{\lambda,n}(c_k, c_{k+1}) + \textit{higher order terms}, \\ \boldsymbol{G}^{total}_{\lambda,n}(c_{k+1}) &= \boldsymbol{G}^{trun.}_{\lambda,\lceil n\rceil}(c_{k+1}) + \boldsymbol{G}^{disc.}_{\lambda,n}(c_{k+1}) + \textit{higher order terms}. \end{aligned}$$

Proof. See Section 3.12. □

The previous theorem links the truncation estimates in Corollary 2.1.1 with the discretization estimates in Corollary 3.4.1 in that their sum controls the total error in the Fer streamers approach to Sturm–Liouville problems.

As can be seen from Theorems 2.1.5 and 3.4.2, the bounds on the local and global discretization errors

$$\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k, c_{k+1}), \qquad \boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_{k+1}),$$

are larger than the bounds on the local and global truncation errors

$$\boldsymbol{L}^{\text{trun.}}_{\lambda,\lceil n\rceil}(c_k, c_{k+1}), \qquad \boldsymbol{G}^{\text{trun.}}_{\lambda,\lceil n\rceil}(c_{k+1}).$$

Hence, according to Theorem 3.4.3, the local and global total errors obey

$$\boldsymbol{L}^{\text{total}}_{\lambda,n}(c_k, c_{k+1}) = \mathcal{O}\left(\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k, c_{k+1})\right), \qquad \boldsymbol{G}^{\text{total}}_{\lambda,n}(c_{k+1}) = \mathcal{O}\left(\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_{k+1})\right).$$

3.5 Numerical results

To illustrate the numerical solution of Sturm–Liouville problems via Fer streamers, with the quadrature schemes in Section 3.3, with global order 4, 7, 10 and 13, uniform over the entire eigenvalue range, consider the Anderssen and de Hoog problem (Anderssen and de Hoog, 1984) defined by:

$$\begin{gathered} a = 0, \quad b = \pi, \quad q(t) = e^t, \\ q_{\min} = q(a), \quad q_{\max} = q(b), \\ y_\lambda(a) - y'_\lambda(a) = 0, \quad y_\lambda(b) + y'_\lambda(b) = 0, \quad \alpha_1 = -\alpha_2 \neq 0, \quad \beta_1 = \beta_2 \neq 0, \end{gathered} \tag{3.5.1}$$

the second Paine problem (Pryce, 1993, p. 281) defined by:

$$\begin{gathered} a = 0, \quad b = \pi, \quad q(t) = \frac{1}{(t + 1/10)^2}, \\ q_{\min} \geq q(b) - 1, \quad q_{\max} = q(a), \\ y_\lambda(a) = y_\lambda(b) = 0, \quad \alpha_1 \neq 0, \quad \beta_1 \neq 0, \quad \alpha_2 = \beta_2 = 0, \end{gathered} \tag{3.5.2}$$

the Coffey–Evans problem (Evans, Coffey and Pryce, 1979; Pryce, 1993, p. 283):

$$\begin{gathered} \beta = 30, \quad a = -\frac{\pi}{2}, \quad b = \frac{\pi}{2}, \quad q(t) = -2\beta\cos(2t) + \beta^2\sin(2t)^2, \\ q_{\min} = -2\beta, \quad q_{\max} = \beta^2 + 1, \\ y_\lambda(a) = y_\lambda(b) = 0, \quad \alpha_1 \neq 0, \quad \beta_1 \neq 0, \quad \alpha_2 = \beta_2 = 0, \end{gathered} \tag{3.5.3}$$

as well as the truncated Gelfand–Levitan problem (Pryce, 1993, p. 283):

$$\begin{gathered} a = 0, \quad b = 100, \quad q(t) = \frac{32\cos(t)(\cos(t) + (2+t)\sin(t))}{(4 + 2t + \sin(2t))^2}, \\ q_{\min} \geq -1, \quad q_{\max} \leq 2, \\ y_\lambda(a) + y'_\lambda(a) = 0, \quad y_\lambda(b) = 0, \quad \alpha_1 = \alpha_2 \neq 0, \quad \beta_1 \neq 0, \quad \beta_2 = 0. \end{gathered} \tag{3.5.4}$$

The numerical results displayed in Figures 3.1–3.2 represent the absolute error and the relative error between an approximation with Fer streamers and one with MATSLISE's package (Ledoux, Daele and Berghe, 2005). To illustrate their power, Fer streamers were generated with the largest possible step size which satisfies Assumption 1.1.1, i.e., with

$$m = \left\lceil (b-a)\sqrt{q_{\max} - q_{\min}} \right\rceil, \qquad h_{\max} = h_{\min} = (b-a)/m,$$

together with

$$n = 1, \qquad n = \log(3)/\log(2), \qquad n = 2, \qquad n = \log(5)/\log(2),$$

in Theorem 3.4.3, i.e., with a method of global order $4, 7, 10, 13$, respectively. It is amazing to observe in Figures 3.1–3.2 that Fer streamers perform well even with extremely large step sizes: in the Anderssen and de Hoog problem with $h_{\max} = h_{\min} = 0.21$, in the second Paine problem with $h_{\max} = h_{\min} = 0.10$, in the Coffey–Evans problem with $h_{\max} = h_{\min} = 0.03$ and in the truncated Gelfand–Levitan with $h_{\max} = h_{\min} = 0.58$. On a related note, it is equally important to observe that the errors in Figures 3.1–3.2 are decreasing with increasing $|\lambda|$, consistent with Theorem 3.4.3. In particular, with Fer streamers with global order $7, 10$, i.e., with $n = \log(3)/\log(2), 2$, the machine precision 10^{-16} is attained for $\lambda \geq 0 \times 10^5$ in the Anderssen and de Hoog problem (3.5.1) and for $\lambda \geq 0.5 \times 10^5$ in the Coffey–Evans problem (3.5.3).

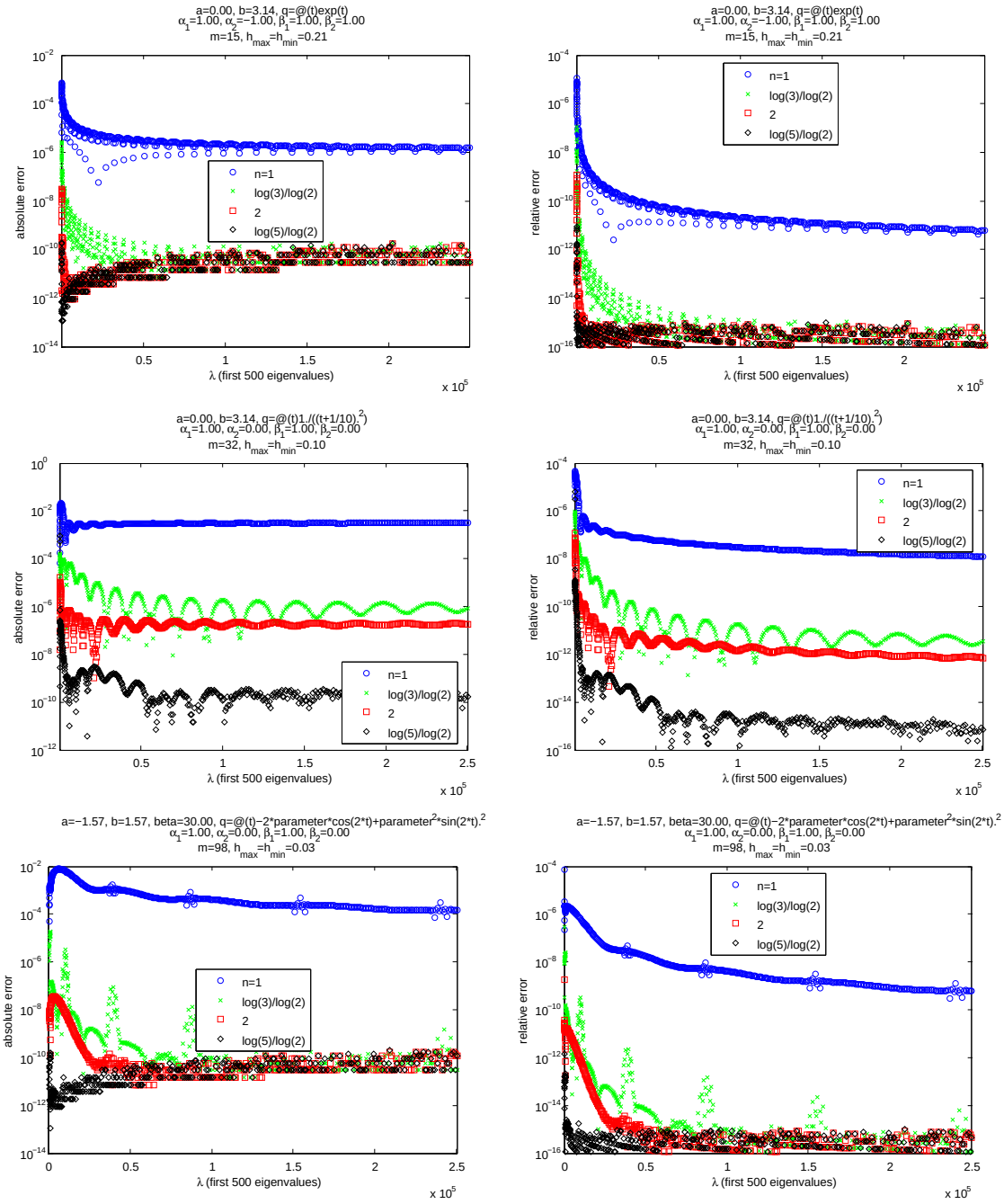


Figure 3.1: Absolute error (left) and relative error (right) with Fer streamers with global order $4, 7, 10, 13$ ($n = 1, \log(3)/\log(2), 2, \log(5)/\log(2)$, respectively) for the Anderssen and de Hoog problem (3.5.1) (top), the second Paine problem (3.5.2) (middle), and the Coffey–Evans problem (3.5.3) (bottom).

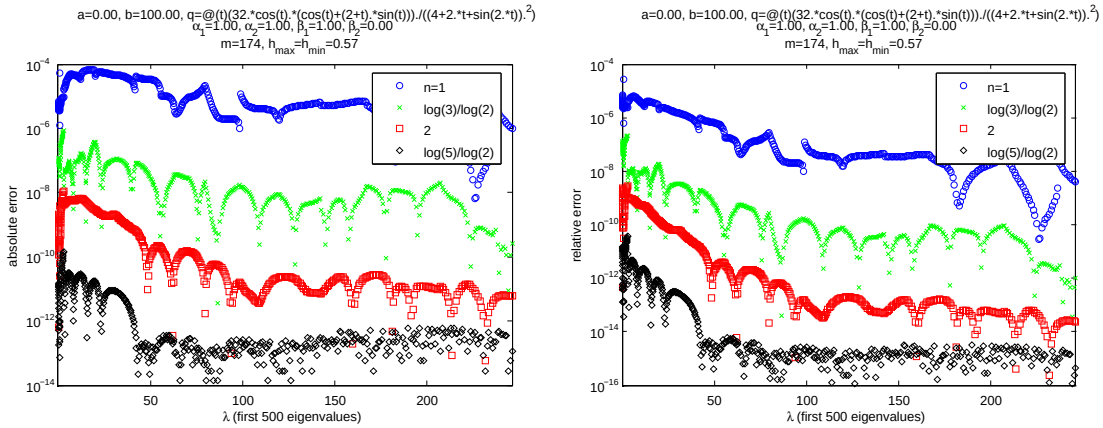


Figure 3.2: Absolute error (left) and relative error (right) with Fer streamers with global order $4, 7, 10, 13$ ($n = 1, \log(3)/\log(2), 2, \log(5)/\log(2)$, respectively) for the truncated Gelfand–Levitan problem (3.5.4).

3.6 Conclusions

It has been shown in this chapter that, in order to preserve the advantageous features of the truncation error in the approximation of $\boldsymbol{Y}_\lambda(c_{k+1})$ by $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$ as obtained in Corollary 2.1.1, also for the discretization error in the approximation of $\tilde{\boldsymbol{Y}}_{\lambda,n}(c_{k+1})$ by $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, while simultaneously minimizing the computational complexity by reducing the number of function evaluations and volume of linear algebra in the discretization schemes, optimal quadrature requires not the simplest representation of each integrand function, but rather relies on an alternative representation carefully designed to comply with a variety of prescribed features. This was the theme of Sections 3.2–3.3, which relied heavily in the closed-form expressions of Fer streamers from Theorem 2.1.3 and, in particular, from Remark 2.1.5.

Tight total error estimates, uniform for every eigenvalue, have also been established in this chapter for the approximation of $\boldsymbol{Y}_\lambda(c_{k+1})$ by $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, that quantify the interplay between the truncation and the discretization in the approach by Fer streamers as well as that guarantee large step sizes uniform over the entire eigenvalue range. This was accomplished in Section 3.4.

Numerical results that illustrate the truncation and discretization of Fer streamers with global orders 4, 7, 10 and 13, have been presented in Section 3.5.

The principal advantage of the Fer streamers approach to Sturm–Liouville problems is that its truncation and discretization error estimates:

(i) hold uniformly for all 'small', 'intermediary' and 'large' eigenvalues, in the sense of (1.1.6)–(1.1.7), and,

(ii) can attain arbitrary high-order.

This is especially significant given that the error estimates in alternative techniques apply only to 'small' or 'large' eigenvalues (c.f. Subsection 1.1.2). Compared with the alternative geometric integration techniques in the right-correction Magnus series (Degani and Schiff, 2006) and in the modified Magnus methods (Ledoux, Daele and Berghe, 2010), which do not possess error estimates uniform over the entire eigenvalue range, the Fer streamers approach presents an interesting trade-off in computational complexity: although it requires an increase in function evaluations for each univariate integral in order to control all 'small', 'intermediary' and 'large' eigenvalues, it also enjoys a significant decrease in linear algebra for each multivariate integral (see Subsection 1.1.5).

To conclude, having derived total error bounds that account for the approximation of $\boldsymbol{Y}_\lambda(c_{k+1})$ by $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$ with the aforementioned advantages of Fer streamers, the question now arises of their implementation aspects, and, in particular, whether one can improve their practical performance by further decreasing the amount of linear algebra in the discretization schemes. Such questions, related to the practical implementation of Fer streamers are investigated in the next Chapter 4, where one can see that their implementation in practice benefits from a reduced Hall basis that leads to a decreased volume of linear algebra in the discretization schemes, which have been realized in a MATLAB package, as reported in the subsequent Chapter 5.

3.7 Proof of Theorem 3.2.1

The representation follows from Remark 2.1.5 and Definition 2.1.4 together with (3.2.4), (3.2.5), (3.2.6), (3.2.7) and (3.2.8). The fact that the derivatives $\zeta^{(j)}_{\lambda,1}(c_k,t)$, $\boldsymbol{U}^{(j)}_{\lambda,1}(c_k,t)$ and $\boldsymbol{V}^{(j)}_{\lambda,1}(c_k,t)$ can be bounded independently of λ follows from the fact that the derivatives of (3.2.9) can be bounded independently of λ.

3.8 Proof of Theorem 3.2.3

The representation follows from Remark 2.1.5 and Definition 2.1.4 together with (3.2.4), (3.2.5), (3.2.6),

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) = 2i(t-c_k)\sqrt{\lambda - \frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}},$$

$$e^{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)} = e^{2i(t-c_k)\left(\sqrt{\lambda-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}-\sqrt{\lambda-q(c_k)}\right)} e^{2i(t-c_k)\sqrt{\lambda-q(c_k)}}.$$

The fact that the derivatives $\zeta_{\lambda,1}^{(j)}(c_k,t)$ and $\boldsymbol{S}_{\lambda,1}^{(j)}(c_k,t)$ can be bounded independently of λ follows from the fact that the derivatives of

$$e^{2i(t-c_k)\left(\sqrt{\lambda-\frac{\int_{c_k}^{t} q(\xi)d\xi}{t-c_k}}-\sqrt{\lambda-q(c_k)}\right)}$$

can be bounded independently of λ.

3.9 Proof of Theorem 3.3.1

Without loss of generality, let $\lambda \in \left[q_{\max}+1, q_{\max}+h_{\max}^{-2}\right]$. The representations proved for this eigenvalue range also hold for $\lambda \in \left[q_{\max}-h_{\max}^{-2}, q_{\min}-1\right]$ because the branch cuts are automatically selected in the various formulæ. The representation follows from Remark 2.1.5 and Definitions 2.1.4 and 3.3.2 together with (3.3.8)–(3.3.13). The terms are arranged in order to make

$$(t-c_k)\begin{bmatrix} t-c_k & (t-c_k)^2 & 1 \end{bmatrix}^{\top}$$

explicit and to make every singularity removable. The fact that the derivatives $\boldsymbol{f}_{\lambda,1}^{(j)}(c_k,t)$ can be bounded independently of λ follows from the fact that the derivatives of (3.3.5), (3.3.6) and (3.3.7) can be bounded independently of λ.

3.10 Proof of Theorem 3.3.3

The representation follows from Remark 2.1.5 and Definitions 2.1.4 and 3.3.2 together with (3.3.8)–(3.3.13). The terms are arranged in order to make

$$(t-c_k)\begin{bmatrix} \frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1 \end{bmatrix}^{\top}$$

explicit and render every singularity removable. The fact that the derivatives $\boldsymbol{g}_{\lambda,1}^{(j)}(c_k,t)$ can be bounded independently of λ follows from the fact that the derivatives of (3.3.5), (3.3.6) and (3.3.7) can be bounded independently of λ.

3.11 Proof of Theorem 3.4.2

The results hold with a proof similar to that of Theorem 2.1.5 for the local and global truncation errors in Definition 2.1.6. Indeed, as with the proof of Theorem 2.1.5, the main obstacle in estimating the local and global discretization errors in Definition 3.4.2, lies in the fact that the lower-left entry of $\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)$ can be arbitrarily large, as

described in Theorem 2.1.4. This main obstacle can be circumvented by calling upon three Baker–Campbell–Hausdorff (BCH) type formulas (2.5.1), (2.5.2) and (2.5.3). For $n = 1$ the local discretization error can be written as

$$\begin{aligned}
\boldsymbol{L}^{\text{disc.}}_{\lambda,1}(c_k, c_{k+1}) &= \log\left(\tilde{\boldsymbol{F}}_{\lambda,\lceil 1\rceil}(c_k, c_{k+1})\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,1}(c_k, c_{k+1})\right) \\
&= \log\left(e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{D}_{\lambda,1}(c_k,c_{k+1})}e^{-\tilde{\boldsymbol{D}}_{\lambda,1,1}(c_k,c_{k+1})}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\right) \\
&= \log\left(e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{E}_{\lambda,1,1}(c_k,c_{k+1})+\text{h.o.t.}}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\right) \\
&= \text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{E}_{\lambda,1,1}(c_k, c_{k+1}) + \text{h.o.t.}\right) \\
&= \text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{E}_{\lambda,1,1}(c_k, c_{k+1})\right) + \text{h.o.t.}
\end{aligned}$$

where the first and second equalities are due to Definitions 2.1.6 and 3.4.2, the third equality is due to (2.5.1) and the fourth equality is due to (2.5.3), whereas for $n \in \{\log(3)/\log(2), 2\}$ the local discretization error can be written as

$$\begin{aligned}
\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k, c_{k+1}) &= \log\left(\tilde{\boldsymbol{F}}_{\lambda,\lceil n\rceil}(c_k, c_{k+1})\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k, c_{k+1})\right) \\
&= \log\Big(e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{D}_{\lambda,1}(c_k,c_{k+1})}e^{\boldsymbol{D}_{\lambda,2}(c_k,c_{k+1})}\times \\
&\qquad \times e^{-\tilde{\boldsymbol{D}}_{\lambda,2,n}(c_k,c_{k+1})}e^{-\tilde{\boldsymbol{D}}_{\lambda,1,n}(c_k,c_{k+1})}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\Big) \\
&= \log\left(e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{E}_{\lambda,1,n}(c_k,c_{k+1})+\boldsymbol{E}_{\lambda,2,n}(c_k,c_{k+1})+\text{h.o.t.}}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\right) \\
&= \log\left(e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{E}_{\lambda,1,n}(c_k,c_{k+1})+\text{h.o.t.}}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\right) \\
&= \text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{E}_{\lambda,1,n}(c_k, c_{k+1}) + \text{h.o.t.}\right) \\
&= \text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{E}_{\lambda,1,n}(c_k, c_{k+1})\right) + \text{h.o.t.}
\end{aligned}$$

where the first and second equalities are due to Definitions 2.1.6 and 3.4.2, the third equality is due to (2.5.1) and the fifth equality is due to (2.5.3). A similar result holds also for $n = \log(5)/\log(2)$. To summarize, for $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$, the local discretization error obeys

$$\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k, c_{k+1}) = \text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{E}_{\lambda,1,n}(c_k, c_{k+1})\right) + \text{h.o.t.}$$

which, together with Theorem 2.1.4 and Theorem 3.4.1, yields the desired estimate. For $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$, the global discretization error obeys the recursion relation with initial condition

$$\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_1) = \boldsymbol{L}^{\text{disc.}}_{\lambda,n}(a, c_1) \tag{3.11.1}$$

and general rule

$$\begin{aligned}
\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_{k+1}) &= \log\left(\tilde{\boldsymbol{Y}}_{\lambda,\lceil n\rceil}(c_{k+1})\tilde{\tilde{\boldsymbol{Y}}}^{-1}_{\lambda,n}(c_{k+1})\right)\\
&= \log\left(\tilde{\boldsymbol{F}}_{\lambda,\lceil n\rceil}(c_k,c_{k+1})\tilde{\boldsymbol{Y}}_{\lambda,\lceil n\rceil}(c_k)\tilde{\tilde{\boldsymbol{Y}}}^{-1}_{\lambda,n}(c_k)\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(\tilde{\boldsymbol{F}}_{\lambda,\lceil n\rceil}(c_k,c_{k+1})e^{\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_k)}\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})}\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k,c_{k+1})e^{\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_k)}\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})}e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}e^{\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_k)+\text{h.o.t.}}e^{-\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})}\exp\left(\text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_k)+\text{h.o.t.}\right)\right)\right)\\
&= \log\left(e^{\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})}\exp\left(\text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_k)\right)+\text{h.o.t.}\right)\right)\\
&= \boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})+\text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_k)\right)+\text{h.o.t.} \qquad (3.11.2)
\end{aligned}$$

where the first, second, third and fourth equalities are due to Definitions 2.1.6 and 3.4.2, the fifth equality is due to (2.5.2), the sixth equality is due to (2.5.3), and the last equality is due to (2.5.1). The global discretization error expressions (3.11.1) and (3.11.2) lead to

$$\begin{aligned}
\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_{k+1}) = {}& \text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))}\left(\boldsymbol{E}_{\lambda,1,n}(c_k,c_{k+1})\right)\\
&+\text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))\exp(\boldsymbol{D}_{\lambda,0}(c_{k-1},c_k))}\left(\boldsymbol{E}_{\lambda,1,n}(c_{k-1},c_k)\right)\\
&+\cdots\\
&+\text{Ad}_{\exp(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1}))\cdots\exp(\boldsymbol{D}_{\lambda,0}(a,c_1))}\left(\boldsymbol{E}_{\lambda,1,n}(a,c_1)\right)\\
&+\text{h.o.t.}
\end{aligned}$$

which, together with Assumption 1.1.1, Theorem 2.1.4 and Theorem 3.4.1, result in the desired estimate.

3.12 Proof of Theorem 3.4.3

The first statement follows from

$$\begin{aligned}
\boldsymbol{L}^{\text{total}}_{\lambda,n}(c_k,c_{k+1}) &= \log\left(\boldsymbol{F}_{\lambda}(c_k,c_{k+1})\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(\boldsymbol{F}_{\lambda}(c_k,c_{k+1})\tilde{\boldsymbol{F}}^{-1}_{\lambda,\lceil n\rceil}(c_k,c_{k+1})\tilde{\boldsymbol{F}}_{\lambda,\lceil n\rceil}(c_k,c_{k+1})\tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k,c_{k+1})\right)\\
&= \log\left(\exp\left(\boldsymbol{L}^{\text{trun.}}_{\lambda,\lceil n\rceil}(c_k,c_{k+1})\right)\exp\left(\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})\right)\right)\\
&= \boldsymbol{L}^{\text{trun.}}_{\lambda,\lceil n\rceil}(c_k,c_{k+1})+\boldsymbol{L}^{\text{disc.}}_{\lambda,n}(c_k,c_{k+1})+\text{higher order terms},
\end{aligned}$$

where the first, second and third equalities are due to Definitions 2.1.6, 3.4.2 and 3.4.3, and the last equality is due to (2.5.1). The second statement follows from

$$\begin{aligned}
\boldsymbol{G}^{\text{total}}_{\lambda,n}(c_{k+1}) &= \log\left(\boldsymbol{Y}_{\lambda}(c_{k+1})\tilde{\tilde{\boldsymbol{Y}}}^{-1}_{\lambda,n}(c_{k+1})\right) \\
&= \log\left(\boldsymbol{Y}_{\lambda}(c_{k+1})\tilde{\boldsymbol{Y}}^{-1}_{\lambda,\lceil n\rceil}(c_{k+1})\tilde{\boldsymbol{Y}}_{\lambda,\lceil n\rceil}(c_{k+1})\tilde{\tilde{\boldsymbol{Y}}}^{-1}_{\lambda,n}(c_{k+1})\right) \\
&= \log\left(\exp\left(\boldsymbol{G}^{\text{trun.}}_{\lambda,\lceil n\rceil}(c_{k+1})\right)\exp\left(\boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_{k+1})\right)\right) \\
&= \boldsymbol{G}^{\text{trun.}}_{\lambda,\lceil n\rceil}(c_{k+1}) + \boldsymbol{G}^{\text{disc.}}_{\lambda,n}(c_{k+1}) + \text{higher order terms}
\end{aligned}$$

where the first, second and third equalities are due to Definitions 2.1.6, 3.4.2 and 3.4.3, and the last equality is due to (2.5.1).

Chapter 4

Decreasing the volume of linear algebra in Fer streamers

Following Fer streamers' truncation and discretization using Lie-algebraic techniques and multivariate oscillatory quadrature in Chapters 2 and 3, the current chapter first recaps these achievements and then discusses Fer streamers' practical implementation with uniform global orders 4, 7, 10 and 13, in the sense of (1.1.6)–(1.1.7), as reported in (Ramos, 2015*b*). In particular, the practical implementation in the present chapter is shown to benefit from a reduced Hall basis which leads to a decreased volume of linear algebra in the given approach.

4.1 A recap of Chapters 2 and 3

Having already an in-depth view of the novel approach to regular Sturm–Liouville problems (1.0.1)–(1.0.2) via Fer streamers in Chapters 2 and 3, the current section summarizes the new set of ideas that surround Fer streamers, necessary for their practical implementation with uniform global orders 4, 7, 10 and 13 with respect to (1.1.6)–(1.1.7). For additional information, including truncation and discretization error bounds, the reader may wish to revisit Chapters 2 and 3.

In a nutshell, the Fer streamers' approach sets out to approximate the exact flow $\boldsymbol{F}_\lambda(c_k, c_{k+1})$ and solution $\boldsymbol{Y}_\lambda(c_{k+1})$ of the initial value problem (1.0.4)–(1.0.5), in Definition 2.1.6. To this end, the approach commences from the well-known Fer expansions integral series summarized in Theorem 2.1.1.

With these integral series in mind, the Fer streamers' approach starts off by using the recursive nature of Fer expansions together with the low-dimensionality of $\mathfrak{sl}(2, \mathbb{R})$ to sum up the infinite sums in Definition 2.1.2 in closed-form, as presented in Theorem 2.1.3. These closed-form expressions, named 'Fer streamers', turn out to be essential to flesh out the magnitude and behaviour of the terms in Fer expansions, required for their integration

in practice. In particular, they yield the representation of the first Fer streamer presented in Remark 2.1.5, which is one of the cornerstones in Chapters 2 and 3.

With the closed-form from Remark 2.1.5 in hand, Definition 3.3.3 introduces the quantity $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$ as a means to expose the fine and coarse scales of $\boldsymbol{B}_{\lambda,1}(c_k, c_k + h_k t)$. These are made precise with Corollaries 3.3.1, 3.3.2 and 3.3.3, which depict the magnitude and behaviour of $\boldsymbol{B}_{\lambda,1}(c_k, c_k + h_k t)$ through $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$. In particular, the magnitude of $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$ is $\mathcal{O}(1)$ with respect to (1.1.6)–(1.1.7) and the behaviour of $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$ changes with $\lambda \in [q_{\max} - h_{\max}^{-2}, +\infty)$ and varies according to: (3.3.1), (3.3.2) and (3.3.3). In particular, $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k \cdot)$ is:

- mildly exponential or oscillatory in (3.3.1) as made clear in Corollary 3.3.1,
- well-behaved in (3.3.2) as made explicit in Corollary 3.3.2,
- mildly to highly oscillatory in (3.3.3) as made explicit in Corollary 3.3.3.

This magnitude and behaviour are important to form an approximation of the quantity $\boldsymbol{B}_{\lambda,1}(c_k, c_k + h_k \cdot)$, which is necessary for the practical implementation of Fer streamers.

With the magnitude and behaviour of $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$ made explicit in Corollaries 3.3.1–3.3.3, Definition 4.1.1 below forms, in line with Subsection 3.3.4, an approximation $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t)$ designed to satisfy two requirements: Firstly, it is such that the difference $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t) - \tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t)$ is uniformly small, with respect to (1.1.6)–(1.1.7). Secondly, it is such that the integrals that appear below in Definition 4.1.2 can be integrated exactly — note the similarity between (3.3.17)–(3.3.21) and (4.1.1)–(4.1.5).

In essence, $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t)$ interpolates the slow varying parts of the term $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$, which, as exposed in Corollaries 3.3.1–3.3.3, depend on (3.3.1)–(3.3.3).

Definition 4.1.1. *Let the interpolation points $\mathcal{T}_{l-1} \subseteq (0, 1]$ be as defined by:*

$$
\begin{aligned}
\mathcal{S}_{11} &:= \{(t+1)/2 : U_{11}(t) = 0\} =: \{u_1, u_2, \ldots, u_{11} : u_1 < u_2 < \cdots < u_{11}\} \subseteq (0, 1) \\
\mathcal{S}_8 &:= \{u_1, u_2, u_4, u_5, u_6, u_8, u_9, u_{10}\} \subseteq \mathcal{S}_{11} \\
\mathcal{S}_5 &:= \{u_2, u_4, u_6, u_8, u_{10}\} = \{(t+1)/2 : U_5(t) = 0\} \subseteq \mathcal{S}_8 \\
\mathcal{S}_2 &:= \{u_4, u_8\} = \{(t+1)/2 : U_2(t) = 0\} \subseteq \mathcal{S}_5 \\
\mathcal{T}_{l-1} &:= \mathcal{S}_{l-2} \cup \{1\}, l \in \{4, 7, 10, 13\}
\end{aligned}
$$

where $U_j(t)$ denotes the j-th Chebyshev polynomial of the second kind. In addition, define also

$$
\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t) \in \mathfrak{sl}(2, \mathbb{R})
$$

in each of (3.3.1), (3.3.2) and (3.3.3) by, respectively:

- *the right hand side of* (3.3.14) *with* $t \mapsto [\boldsymbol{f}_{\lambda,1}(c_k, c_k + h_k t)]_{j,1}$ *replaced by polynomial interpolation at* $\mathcal{T}_{l-1}$,
- *the right hand side of* (3.3.15) *with* $t \mapsto [\boldsymbol{\iota}_{\lambda,1}(c_k, c_k + h_k t)]_{j,1}$ *replaced by polynomial interpolation at* $\mathcal{T}_{l-1}$,
- *the right hand side of* (3.3.16) *with* $t \mapsto [\boldsymbol{g}_{\lambda,1}(c_k, c_k + h_k t)]_{j,1}$ *replaced by polynomial interpolation at* $\mathcal{T}_{l-1}$.

With the machinery introduced above, Definition 4.1.2 below sets out the key elements $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k, c_{k+1})$ and $\tilde{\boldsymbol{I}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k, c_{k+1})$, which emerge in the truncated and discretized flow and solution that appear at the end of this section in Definition 4.1.3 and Theorem 4.1.1. In particular, $\tilde{\boldsymbol{I}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k, c_{k+1})$ are given by a rescaling of $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k, c_{k+1})$, which, by construction, can be integrated exactly.

Definition 4.1.2. *Let*

$$\tilde{\boldsymbol{I}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_{k+1}) := \int_0^1 \tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t)dt, \tag{4.1.1}$$

$$\tilde{\boldsymbol{I}}^{fine}_{\lambda,2,\mathcal{T}_{l-1}}(c_k, c_{k+1}) := \int_0^1 \int_0^{t_1} \big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_2), \tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_1)\big] dt_2 dt_1, \tag{4.1.2}$$

$$\tilde{\boldsymbol{I}}^{fine}_{\lambda,3,\mathcal{T}_{l-1}}(c_k, c_{k+1}) := \int_0^1 \int_0^{t_1} \int_0^{t_1} \big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_3), \big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_2), \tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_1)\big]\big] dt_3 dt_2 dt_1, \tag{4.1.3}$$

$$\tilde{\boldsymbol{I}}^{fine}_{\lambda,4,\mathcal{T}_{l-1}}(c_k, c_{k+1}) := \int_0^1 \int_0^{t_1} \int_0^{t_1} \int_0^{t_1} \big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_4), \big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_3), \big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_2), \tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_1)\big]\big]\big] dt_4 dt_3 dt_2 dt_1, \tag{4.1.4}$$

$$\tilde{\boldsymbol{I}}^{fine}_{\lambda,5,\mathcal{T}_{l-1}}(c_k, c_{k+1}) := \int_0^1 \int_0^{t_1} \int_0^{t_1} \int_0^{t_2} \big[\big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_4), \tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_2)\big], \big[\tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_3), \tilde{\boldsymbol{B}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t_1)\big]\big] dt_4 dt_3 dt_2 dt_1. \tag{4.1.5}$$

In addition, let also $\tilde{\boldsymbol{I}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k,c_{k+1}),\ldots,\tilde{\boldsymbol{I}}_{\lambda,5,\mathcal{T}_{l-1}}(c_k,c_{k+1})\in\mathfrak{sl}(2,\mathbb{R})$ *be the unique elements which satisfy*

$$\begin{aligned}
\boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big) &:= \boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}^{fine}_{\lambda,1,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big)\\
&\quad\odot\begin{cases} h_k^2\begin{bmatrix} h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ h_k^2\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2},\end{cases}\\
\boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}_{\lambda,2,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big) &:= \boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}^{fine}_{\lambda,2,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big)\\
&\quad\odot\begin{cases} h_k^5\begin{bmatrix} h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ \frac{h_k^4}{\sqrt{\lambda-q\,(c_k)}}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2},\end{cases}\\
\boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}_{\lambda,3,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big) &:= \boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}^{fine}_{\lambda,3,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big)\\
&\quad\odot\begin{cases} h_k^8\begin{bmatrix} h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ \frac{h_k^6}{\lambda-q\,(c_k)}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2},\end{cases}\\
\boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}_{\lambda,4,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big) &:= \boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}^{fine}_{\lambda,4,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big)\\
&\quad\odot\begin{cases} h_k^{11}\begin{bmatrix} h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ \frac{h_k^8}{(\lambda-q\,(c_k))^{\frac{3}{2}}}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2},\end{cases}\\
\boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}_{\lambda,5,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big) &:= \boldsymbol{\pi}\big(\tilde{\boldsymbol{I}}^{fine}_{\lambda,5,\mathcal{T}_{l-1}}(c_k,c_{k+1})\big)\\
&\quad\odot\begin{cases} h_k^{11}\begin{bmatrix} h_k & h_k^2 & 1\end{bmatrix}^\top, & |\lambda-q_{max}|\leq h_{max}^{-2},\\ \frac{h_k^8}{(\lambda-q\,(c_k))^{\frac{3}{2}}}\begin{bmatrix}\frac{1}{\sqrt{\lambda-q(c_k)}} & \frac{1}{\lambda-q(c_k)} & 1\end{bmatrix}^\top, & \lambda-q_{max}\geq h_{max}^{-2}.\end{cases}
\end{aligned}$$

In order to approximate the exact flow $\boldsymbol{F}_\lambda(c_k,c_{k+1})$ and solution $\boldsymbol{Y}_\lambda(c_{k+1})$ of the initial value problem (1.0.4)–(1.0.5) in Definition 2.1.6, Definition 4.1.3 below reformulates the truncated and discretized flow $\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k,c_{k+1})$ and solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, together with the local $\boldsymbol{L}^{\text{total}}_{\lambda,n}(c_k,c_{k+1})$ and global $\boldsymbol{G}^{\text{total}}_{\lambda,n}(c_{k+1})$ errors that characterize each approximation, found in previous chapters. In particular, in line with Subsection 1.1.4, $\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k,c_{k+1})$ and $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$ belong to the Lie group $\mathrm{SL}(2,\mathbb{R})$, whereas $\boldsymbol{L}^{\text{total}}_{\lambda,n}(c_k,c_{k+1})$ and $\boldsymbol{G}^{\text{total}}_{\lambda,n}(c_{k+1})$ lie in the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$.

As established above in Chapters 2 and 3, and testified again in Theorem 4.1.1 below,

Fer streamers then achieve global order $g \in \{4, 7, 10, 13\}$, with respect to (1.1.6)–(1.1.7), by approximating $\boldsymbol{Y}_\lambda(c_{k+1})$ with $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$ where $n = \log((g+2)/3)/\log(2)$.

Definition 4.1.3 (Reformulated from Definitions 3.4.2 and 3.4.3)**.** *If* $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$*, define*

$$
\begin{aligned}
\tilde{\tilde{\boldsymbol{F}}}_{\lambda,1}(c_k, c_{k+1}) &:= e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{I}}_{\lambda,1,\mathcal{T}_3}(c_k,c_{k+1})}, \\
\tilde{\tilde{\boldsymbol{F}}}_{\lambda,\log(3)/\log(2)}(c_k, c_{k+1}) &:= e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{I}}_{\lambda,1,\mathcal{T}_6}(c_k,c_{k+1})} \\
&\quad \times e^{-\frac{1}{2}\tilde{\boldsymbol{I}}_{\lambda,2,\mathcal{T}_3}(c_k,c_{k+1})}, \\
\tilde{\tilde{\boldsymbol{F}}}_{\lambda,2}(c_k, c_{k+1}) &:= e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{I}}_{\lambda,1,\mathcal{T}_9}(c_k,c_{k+1})} \\
&\quad \times e^{-\frac{1}{2}\tilde{\boldsymbol{I}}_{\lambda,2,\mathcal{T}_6}(c_k,c_{k+1})+\frac{1}{3}\tilde{\boldsymbol{I}}_{\lambda,3,\mathcal{T}_3}(c_k,c_{k+1})}, \\
\tilde{\tilde{\boldsymbol{F}}}_{\lambda,\log(5)/\log(2)}(c_k, c_{k+1}) &:= e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} e^{\tilde{\boldsymbol{I}}_{\lambda,1,\mathcal{T}_{12}}(c_k,c_{k+1})} \\
&\quad \times e^{-\frac{1}{2}\tilde{\boldsymbol{I}}_{\lambda,2,\mathcal{T}_9}(c_k,c_{k+1})+\frac{1}{3}\tilde{\boldsymbol{I}}_{\lambda,3,\mathcal{T}_6}(c_k,c_{k+1})-\frac{1}{8}\tilde{\boldsymbol{I}}_{\lambda,4,\mathcal{T}_3}(c_k,c_{k+1})} \\
&\quad \times e^{-\frac{1}{8}\tilde{\boldsymbol{I}}_{\lambda,5,\mathcal{T}_3}(c_k,c_{k+1})}, \\
\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1}) &:= \tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1}) \cdots \tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_1, c_2) \tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(a, c_1), \\
\boldsymbol{L}^{total}_{\lambda,n}(c_k, c_{k+1}) &:= \log\left(\boldsymbol{F}_\lambda(c_k, c_{k+1}) \tilde{\tilde{\boldsymbol{F}}}^{-1}_{\lambda,n}(c_k, c_{k+1})\right), \\
\boldsymbol{G}^{total}_{\lambda,n}(c_{k+1}) &:= \log\left(\boldsymbol{Y}_\lambda(c_{k+1}) \tilde{\tilde{\boldsymbol{Y}}}^{-1}_{\lambda,n}(c_{k+1})\right).
\end{aligned}
$$

Theorem 4.1.1 (Ramos, 2015*b*)**.** *If* $n \in \{1, \log(3)/\log(2), 2, \log(5)/\log(2)\}$*, and* (3.3.1)*,* (3.3.2) *or* (3.3.3)*, then, in the uniform regime* (1.1.6)–(1.1.7)*,*

$$
\boldsymbol{\pi}\left(\boldsymbol{L}^{total}_{\lambda,n}(c_k, c_{k+1})\right) = h_{\max}^{3\times 2^n - 1} \begin{cases} \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}\left(h^2_{\max}\right) & \mathcal{O}(1) \end{bmatrix}^\top, & |\lambda - q_{max}| \leq h_{max}^{-2}, \\ \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} & \mathcal{O}(1) \end{bmatrix}^\top, & \lambda - q_{max} \geq h_{max}^{-2}, \end{cases}
$$

$$
\boldsymbol{\pi}\left(\boldsymbol{G}^{total}_{\lambda,n}(c_{k+1})\right) = h_{\max}^{3\times 2^n - 2} \begin{cases} \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}\left(h^2_{\max}\right) & \mathcal{O}(1) \end{bmatrix}^\top, & |\lambda - q_{max}| \leq h_{max}^{-2}, \\ \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} & \mathcal{O}(1) \end{bmatrix}^\top, & \lambda - q_{max} \geq h_{max}^{-2}. \end{cases}
$$

Proof. Follows from the total error bounds in Theorem 3.4.3, together with the truncation error bounds in Theorem 2.1.5 and the discretization error bounds in Theorem 3.4.2. □

As indicated in Definition 4.1.3 and Theorem 4.1.1, the Fer streamers approach with global order $g \in \{4, 7, 10, 13\}$, uses the polynomial interpolation in Definition 4.1.1 to approximate $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$ with $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t)$, $l \in \{4, 7, 10, 13\} \cap [0, g]$ uniformly

with respect to (1.1.6)–(1.1.7), which requires the data (c.f., Subsubsection 3.3.4.3):

$$\{q(a)\} \cup \left(\bigcup_{k=0}^{m-1} \left\{ q(c_k + h_k t), \int_{c_k}^{c_k+h_k t} q(\xi) d\xi : t \in \mathcal{T}_{l-1} \right\} \right).$$

Since the antiderivative of the potential is usually unavailable in closed-form, one approximates, up to local order, the antiderivative data

$$\left\{ \int_{c_k}^{c_k+h_k t} q(\xi) d\xi : t \in \mathcal{T}_{l-1} \right\}$$

by the polynomial interpolation of $q(\xi)$ in $\xi \in [c_k, c_{k+1}]$ with the potential data

$$\{q(c_k)\} \cup \{q(c_k + h_k t) : t \in \mathcal{T}_{l-1}\}$$

and the exact integration of the result. Since $\mathcal{T}_3 \subseteq \mathcal{T}_6 \subseteq \mathcal{T}_9 \subseteq \mathcal{T}_{12}$ (c.f. Definition 4.1.1), to attain global order $p+1 \in \{4, 7, 10, 13\}$, Fer streamers evaluate $q(a)$ and $q(c_k + h_k \cdot)$, $k \in \{0, \ldots, m-1\}$, at the p points in $\mathcal{T}_p$, in accordance with Subsection 1.1.3.

4.2 Practical implementation of Fer streamers

Following the description of the uniform approximations provided by Fer streamers for the solution of the initial value problem (1.0.4)–(1.0.5) in the previous section, the present section now bridges between theoretical construction and practical implementation. In particular, Subsections 4.2.1–4.2.2 below examine the practical implementation of the truncated and discretized solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, whereas Sections 5.1–5.3 discuss the use of the computed data $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$ to approximate the eigenvalues and eigenfunctions of the boundary value problem (1.0.1)–(1.0.2).

4.2.1 Reduced Hall basis for Fer streamers

In view of Definitions 4.1.2 and 4.1.3, the implementation of the truncated and discretized solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$ boils down to the computation of $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k, c_{k+1})$. More precisely, according to Theorem 4.1.1, to attain global order $g \in \{4, 7, 10, 13\}$, in the sense of (1.1.6)–(1.1.7), one may approximate $\boldsymbol{Y}_\lambda(c_{k+1})$ by $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$ with $n = \log((g+2)/3)/\log(2)$, the implementation of which reduces to the computation of:

- $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_3}(c_k, c_{k+1})$, for $g = 4$,
- $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_6}(c_k, c_{k+1})$, $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,2,\mathcal{T}_3}(c_k, c_{k+1})$, for $g = 7$,
- $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_9}(c_k, c_{k+1})$, $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,2,\mathcal{T}_6}(c_k, c_{k+1})$, $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,3,\mathcal{T}_3}(c_k, c_{k+1})$, for $g = 10$,

- $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{12}}(c_k,c_{k+1})$, $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,2,\mathcal{T}_9}(c_k,c_{k+1})$, $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,3,\mathcal{T}_6}(c_k,c_{k+1})$, $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,4,\mathcal{T}_3}(c_k,c_{k+1})$, $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,5,\mathcal{T}_3}(c_k,c_{k+1})$, for $g=13$.

As discussed in this subsection, since, by construction, each integral (4.1.1)–(4.1.5) can be integrated exactly, the question then becomes how to achieve such computation with the least volume of linear algebra and, by extension, computational time.

To minimize the length of this subsection while retaining its essential message, the discussion focuses on the reduction of the volume of linear algebra for the terms with three interpolation points across global orders 4, 7, 10 and 13:

$$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_3}(c_k,c_{k+1}),\quad \tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,2,\mathcal{T}_3}(c_k,c_{k+1}),\quad \tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,3,\mathcal{T}_3}(c_k,c_{k+1}),\quad \tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,4,\mathcal{T}_3}(c_k,c_{k+1}),\quad \tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,5,\mathcal{T}_3}(c_k,c_{k+1}), \tag{4.2.1}$$

which reveal the ins and outs also for the computation of the other terms.

Recalling the construction of $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k,c_k+h_kt)$ in Definition 4.1.1, it is clear that, for each fixed numerical mesh, the computation of (4.1.1)–(4.1.5), depends on which interval (3.3.1), (3.3.2) or (3.3.3), λ lies on. As an example, with three interpolation points, by solving a linear system exactly, one can write:

$$\begin{aligned}
&\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_3}(c_k,c_k+h_kt)\\
&\left\{\begin{aligned}
&=: \frac{1-\cos\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)}{\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)^2}t^2\left(t^2\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3}+\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,3,\mathcal{T}_3}\right)\\
&\quad+\cos\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)t\Big(t^4\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t^3\mathscr{D}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3}\\
&\qquad\qquad+t^2\mathscr{G}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t\mathscr{E}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+\mathscr{C}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3}\Big)\\
&\quad+\frac{\sin\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)}{\omega_{\lambda,1}(c_k,c_{k+1})t}t\left(t^3\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,4,\mathcal{T}_3}+t^2\mathscr{E}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3}+t\mathscr{E}^{\boldsymbol{f},c_k,h_k}_{\lambda,3,\mathcal{T}_3}+\mathscr{C}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3}\right)\\
&\quad+\phi\left(i\cdot\omega_{\lambda,1}(c_k,c_{k+1})t\right)t^3\left(t^2\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3}+t\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,3,\mathcal{T}_3}+\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,4,\mathcal{T}_3}\right), &&\Leftarrow (3.3.1),\\
&=: t\left(t^4\mathscr{B}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t^3\mathscr{D}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t^2\mathscr{G}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t\mathscr{E}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+\mathscr{C}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}\right), &&\Leftarrow (3.3.2),\\
&=: \frac{1-\cos\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)}{\omega_{\lambda,1}(c_k,c_{k+1})t}t\left(t^2\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,2,\mathcal{T}_3}+\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,3,\mathcal{T}_3}\right)\\
&\quad+\cos\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)t\left(t^2\mathscr{G}^{\boldsymbol{g},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t\mathscr{G}^{\boldsymbol{g},c_k,h_k}_{\lambda,2,\mathcal{T}_3}+\mathscr{G}^{\boldsymbol{g},c_k,h_k}_{\lambda,3,\mathcal{T}_3}\right)\\
&\quad+\frac{\sin\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)}{\omega_{\lambda,1}(c_k,c_{k+1})t}t\left(t^2\mathscr{F}^{\boldsymbol{g},c_k,h_k}_{\lambda,1,\mathcal{T}_3}+t\mathscr{F}^{\boldsymbol{g},c_k,h_k}_{\lambda,2,\mathcal{T}_3}+\mathscr{F}^{\boldsymbol{g},c_k,h_k}_{\lambda,3,\mathcal{T}_3}\right)\\
&\quad+\sin\left(\omega_{\lambda,1}(c_k,c_{k+1})t\right)t\left(t^2\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,4,\mathcal{T}_3}+t\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,5,\mathcal{T}_3}+\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,6,\mathcal{T}_3}\right), &&\Leftarrow (3.3.3),
\end{aligned}\right.
\end{aligned}$$

where each group of matrices

$$\begin{aligned}
&\Big\{\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,3,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{f},c_k,h_k}_{\lambda,4,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\\
&\quad\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,2,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,3,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,4,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,5,\mathcal{T}_3},\mathscr{A}^{\boldsymbol{g},c_k,h_k}_{\lambda,6,\mathcal{T}_3}\Big\},
\end{aligned} \tag{4.2.2}$$

$$\left\{\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3},\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,3,\mathcal{T}_3},\mathscr{B}^{\boldsymbol{f},c_k,h_k}_{\lambda,4,\mathcal{T}_3},\mathscr{B}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}\right\}, \tag{4.2.3}$$

$$\left\{\mathscr{C}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{C}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3},\mathscr{C}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}\right\}, \tag{4.2.4}$$

$$\left\{\mathscr{D}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{D}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}\right\}, \tag{4.2.5}$$

$$\left\{\mathscr{E}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{E}^{\boldsymbol{f},c_k,h_k}_{\lambda,2,\mathcal{T}_3},\mathscr{E}^{\boldsymbol{f},c_k,h_k}_{\lambda,3,\mathcal{T}_3},\mathscr{E}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3}\right\}, \tag{4.2.6}$$

$$\left\{\mathscr{F}^{\boldsymbol{g},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{F}^{\boldsymbol{g},c_k,h_k}_{\lambda,2,\mathcal{T}_3},\mathscr{F}^{\boldsymbol{g},c_k,h_k}_{\lambda,3,\mathcal{T}_3}\right\}, \tag{4.2.7}$$

$$\left\{\mathscr{G}^{\boldsymbol{f},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{G}^{\boldsymbol{\iota},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{G}^{\boldsymbol{g},c_k,h_k}_{\lambda,1,\mathcal{T}_3},\mathscr{G}^{\boldsymbol{g},c_k,h_k}_{\lambda,2,\mathcal{T}_3},\mathscr{G}^{\boldsymbol{g},c_k,h_k}_{\lambda,3,\mathcal{T}_3}\right\}, \tag{4.2.8}$$

possesses a certain structure. More concretely, if

$$\boldsymbol{E}_1 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \boldsymbol{E}_2 := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \boldsymbol{E}_3 := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

then (4.2.2)–(4.2.8) exhibit the following features:

$$\mathscr{A} \in \operatorname{span}\{\boldsymbol{E}_1\}, \qquad \mathscr{D} \in \operatorname{span}\{\boldsymbol{E}_1,\boldsymbol{E}_2\}, \tag{4.2.9}$$

$$\mathscr{B} \in \operatorname{span}\{\boldsymbol{E}_2\}, \qquad \mathscr{E} \in \operatorname{span}\{\boldsymbol{E}_1,\boldsymbol{E}_3\}, \tag{4.2.10}$$

$$\mathscr{C} \in \operatorname{span}\{\boldsymbol{E}_3\}, \qquad \mathscr{F} \in \operatorname{span}\{\boldsymbol{E}_2,\boldsymbol{E}_3\}, \qquad \mathscr{G} \in \operatorname{span}\{\boldsymbol{E}_1,\boldsymbol{E}_2,\boldsymbol{E}_3\}. \tag{4.2.11}$$

Even though, as said before, the presentation focuses on the decrease of the volume of linear algebra for the exact integration of (4.2.1) with three interpolation points, it is pertinent at this moment to inform the reader that the aforementioned representation of $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k+h_k t)$ in terms of matrices of type $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$, $\mathscr{E}$, $\mathscr{F}$ and $\mathscr{G}$, made explicit above for three interpolation points, holds similarly for any number of points, given that, in general, by solving a linear system exactly, one can write $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{l-1}}(c_k, c_k + h_k t)$ as a linear combination of:

- $((l-1)+1)$ $\mathscr{A}$'s, $((l-1)+1)$ $\mathscr{B}$'s, 2 $\mathscr{C}$'s, 1 $\mathscr{D}$, $(l-1)$ $\mathscr{E}$'s and $((l-1)-2)$ $\mathscr{G}$'s, in (3.3.1),
- 1 $\mathscr{B}$, 1 $\mathscr{C}$, 1 $\mathscr{D}$, 1 $\mathscr{E}$ and $((l-1)-2)$ $\mathscr{G}$'s, in (3.3.2),
- $2(l-1)$ $\mathscr{A}$'s, $(l-1)$ $\mathscr{F}$'s and $(l-1)$ $\mathscr{G}$'s, in (3.3.3).

With this in mind, returning to the exact integration of (4.2.1), it is now convenient to aggregate the matrices in (4.2.2)–(4.2.8) according to each eigenvalue range (3.3.1), (3.3.2) or (3.3.3), and to denote them more generically by:

$$\begin{aligned}\Big\{&\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,1},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,2},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,3},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,4},\\ &\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,5},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,6},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,7},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,8},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,9},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,10},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,11},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,12},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,13},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,14},\boldsymbol{Z}^{\boldsymbol{f}}_{\lambda,15}\Big\},\end{aligned} \tag{4.2.12}$$

$$\left\{\boldsymbol{Z}^{\boldsymbol{\iota}}_{\lambda,1},\boldsymbol{Z}^{\boldsymbol{\iota}}_{\lambda,2},\boldsymbol{Z}^{\boldsymbol{\iota}}_{\lambda,3},\boldsymbol{Z}^{\boldsymbol{\iota}}_{\lambda,4},\boldsymbol{Z}^{\boldsymbol{\iota}}_{\lambda,5}\right\}, \tag{4.2.13}$$

$$\left\{\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,1},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,2},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,3},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,4},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,5},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,6},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,7},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,8},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,9},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,10},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,11},\boldsymbol{Z}^{\boldsymbol{g}}_{\lambda,12}\right\}. \tag{4.2.14}$$

Gauging upon the definition of (4.2.1) in (4.1.1)–(4.1.5), noting that the integrands of (4.2.1) are, respectively, 0, 1, 2, 3 and 3 commutators between $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_3}(c_k, c_k + h_k\xi)$ evaluated at various ξ, while at the same time recalling that $\tilde{\boldsymbol{B}}^{\text{fine}}_{\lambda,1,\mathcal{T}_3}(c_k, c_k + h_k\xi)$ has been given above as a linear combination of $\boldsymbol{Z}$'s, where each scalar coefficient of each matrix $\boldsymbol{Z}$ is a function of ξ, it becomes clear that to integrate each (4.2.1), one must expand each commutator representation of each integrand via the bilinear properties of the commutator

$$\begin{aligned}
\big[\boldsymbol{Z}_{\lambda,j_1} + \boldsymbol{Z}_{\lambda,j_2}, \boldsymbol{Z}_{\lambda,j_3}\big] &= \big[\boldsymbol{Z}_{\lambda,j_1}, \boldsymbol{Z}_{\lambda,j_3}\big] + \big[\boldsymbol{Z}_{\lambda,j_2}, \boldsymbol{Z}_{\lambda,j_3}\big],\\
\big[\boldsymbol{Z}_{\lambda,j_1}, \boldsymbol{Z}_{\lambda,j_2} + \boldsymbol{Z}_{\lambda,j_3}\big] &= \big[\boldsymbol{Z}_{\lambda,j_1}, \boldsymbol{Z}_{\lambda,j_2}\big] + \big[\boldsymbol{Z}_{\lambda,j_1}, \boldsymbol{Z}_{\lambda,j_3}\big],\\
\big[c\boldsymbol{Z}_{\lambda,j_1}, \boldsymbol{Z}_{\lambda,j_2}\big] &= c\big[\boldsymbol{Z}_{\lambda,j_1}, \boldsymbol{Z}_{\lambda,j_2}\big], \quad c \in \mathbb{R},\\
\big[\boldsymbol{Z}_{\lambda,j_1}, c\boldsymbol{Z}_{\lambda,j_2}\big] &= c\big[\boldsymbol{Z}_{\lambda,j_1}, \boldsymbol{Z}_{\lambda,j_2}\big], \quad c \in \mathbb{R},
\end{aligned}$$

to single out the scalar functions that require integration. This, of course, represents each integrand of (4.2.1) as a linear combination of the elements with, respectively, 0, 1, 2, 3 and 3 commutators in the free magma of each alphabet (4.2.12), (4.2.13) or (4.2.14), the size of which grows significantly with the size of the alphabet and the number of commutators, as depicted in Table 4.1 under "free magma".

The volume of linear algebra mentioned above then relates to the number of commutators that result from such procedure.

Fortunately, there exist three mechanisms that can be used to decrease this volume of linear algebra. These are:

- Firstly, Free Lie algebra (FLA) techniques and Hall basis, which lead to fewer commutators via a systematic use of commutator identities such as: skew symmetry, Jacobi's identity, etc,
- Secondly, when collected in a Hall basis (which varies with the ordering of the alphabet), certain linear combinations between different integrands are then identically zero,
- Thirdly, when collected in a Hall basis (which depends on the ordering of the alphabet), certain linear combinations between different integrands integrate exactly to zero.

The remainder of this section then concerns a brief description of the savings achieved via these mechanisms, the first of which is well-known in the literature, whereas the second and third arise now from the practical implementation of Fer streamers.

The first mechanism above is well-known, an excellent reference being (Munthe–Kaas and Owren, 1999). In a nutshell, FLA techniques and Hall basis, diminish the number of commutators by judiciously invoking skew symmetry, Jacobi's identity and other relations:

$$\begin{aligned}
\big[\boldsymbol{Z}_{\lambda,j_1},\boldsymbol{Z}_{\lambda,j_2}\big] &= -\big[\boldsymbol{Z}_{\lambda,j_2},\boldsymbol{Z}_{\lambda,j_1}\big], && (4.2.15)\\
0 &= \big[\boldsymbol{Z}_{\lambda,j_1},\big[\boldsymbol{Z}_{\lambda,j_2},\boldsymbol{Z}_{\lambda,j_3}\big]\big]\\
&\quad+\big[\boldsymbol{Z}_{\lambda,j_2},\big[\boldsymbol{Z}_{\lambda,j_3},\boldsymbol{Z}_{\lambda,j_1}\big]\big]\\
&\quad+\big[\boldsymbol{Z}_{\lambda,j_3},\big[\boldsymbol{Z}_{\lambda,j_1},\boldsymbol{Z}_{\lambda,j_2}\big]\big], && (4.2.16)\\
\big[\big[\boldsymbol{Z}_{\lambda,j_1},\boldsymbol{Z}_{\lambda,j_3}\big],\big[\boldsymbol{Z}_{\lambda,j_2},\boldsymbol{Z}_{\lambda,j_4}\big]\big] &= -\big[\boldsymbol{Z}_{\lambda,j_4},\big[\boldsymbol{Z}_{\lambda,j_3},\big[\boldsymbol{Z}_{\lambda,j_2},\boldsymbol{Z}_{\lambda,j_1}\big]\big]\big]\\
&\quad-\big[\boldsymbol{Z}_{\lambda,j_1},\big[\boldsymbol{Z}_{\lambda,j_4},\big[\boldsymbol{Z}_{\lambda,j_3},\boldsymbol{Z}_{\lambda,j_2}\big]\big]\big]\\
&\quad-\big[\boldsymbol{Z}_{\lambda,j_2},\big[\boldsymbol{Z}_{\lambda,j_1},\big[\boldsymbol{Z}_{\lambda,j_4},\boldsymbol{Z}_{\lambda,j_3}\big]\big]\big]\\
&\quad-\big[\boldsymbol{Z}_{\lambda,j_3},\big[\boldsymbol{Z}_{\lambda,j_2},\big[\boldsymbol{Z}_{\lambda,j_1},\boldsymbol{Z}_{\lambda,j_4}\big]\big]\big], && (4.2.17)
\end{aligned}$$

to remove commutators that become redundant in light of such equalities. As illustrated in Table 4.1 under "Hall basis", the number of terms decreases substantially from the free magma to the Hall basis. On this point, it is important to note that the MATLAB package DiffMan (Engø, Marthinsen and Munthe–Kaas, 1999) has been used to confirm the coefficient expansion of the various terms in the Hall basis, which have been used to sort up the data required for global order up to 13.

The second mechanism above originates from the observation that while it is unquestionable that expressing the commutators in a Hall basis results in a significant decrease of the volume of linear algebra, it is equally true that, by construction, a Hall basis does not take into account any structure that each letter '$\boldsymbol{Z}$' of each alphabet (4.2.12), (4.2.13) or (4.2.14) might possess, creating a chance for further reduction. In keeping with this train of thought, in view of (4.2.9)–(4.2.11), one realizes that the equalities

$$\big[\boldsymbol{E}_1,\boldsymbol{E}_2\big] = 2\boldsymbol{E}_2, \qquad \big[\boldsymbol{E}_1,\boldsymbol{E}_3\big] = -2\boldsymbol{E}_3, \qquad \big[\boldsymbol{E}_2,\boldsymbol{E}_3\big] = \boldsymbol{E}_1,$$

give rise to many relations that are not captured by a Hall basis, such as, for instance:

$$\begin{aligned}
[\mathscr{A}_{\lambda,j_1},\mathscr{A}_{\lambda,j_2}] &= \mathbf{0},\forall j_1,\forall j_2,\\
[[\mathscr{A}_{\lambda,j_1},\mathscr{G}_{\lambda,j}],[\mathscr{A}_{\lambda,j_2},\mathscr{G}_{\lambda,j}]] &= \mathbf{0},\forall j_1,\forall j_2,\forall j,\\
[\mathscr{A}_{\lambda,j_1},[\mathscr{F}_{\lambda,j_2},\mathscr{F}_{\lambda,j_3}]] &= \mathbf{0},\forall j_1,\forall j_2,\forall j_3,
\end{aligned}$$

which, with a slight abuse of notation, can be summarized more simply as:

$$[\mathscr{A},\mathscr{A}] = \mathbf{0}, \qquad [[\mathscr{A},\mathscr{G}_{\lambda,j}],[\mathscr{A},\mathscr{G}_{\lambda,j}]] = \mathbf{0}, \qquad [\mathscr{A},[\mathscr{F},\mathscr{F}]] = \mathbf{0}.$$

This abuse allows to write all relations not captured by a Hall basis up to 3 commutators:

$$[\mathscr{A},\mathscr{A}]=\mathbf{0}, \tag{4.2.18}$$

$$[\mathscr{B},\mathscr{B}]=\mathbf{0}, \qquad [\mathscr{C},\mathscr{C}]=\mathbf{0}, \tag{4.2.19}$$

$$[\mathscr{A},[\mathscr{B},\mathscr{C}]]=\mathbf{0}, \qquad [\mathscr{C},[\mathscr{E},\mathscr{E}]]=\mathbf{0}, \tag{4.2.20}$$

$$[\mathscr{B},[\mathscr{A},\mathscr{B}]]=\mathbf{0}, \qquad [\mathscr{C},[\mathscr{A},\mathscr{C}]]=\mathbf{0}, \tag{4.2.21}$$

$$[\mathscr{B},[\mathscr{A},\mathscr{D}]]=\mathbf{0}, \qquad [\mathscr{C},[\mathscr{A},\mathscr{E}]]=\mathbf{0}, \tag{4.2.22}$$

$$[\mathscr{B},[\mathscr{B},\mathscr{D}]]=\mathbf{0}, \qquad [\mathscr{C},[\mathscr{C},\mathscr{E}]]=\mathbf{0}, \tag{4.2.23}$$

$$[\mathscr{A},[\mathscr{B},[\mathscr{A},\mathscr{C}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{A},[\mathscr{B},\mathscr{G}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{A},[\mathscr{C},\mathscr{E}]]]=\mathbf{0}, \tag{4.2.24}$$

$$[\mathscr{A},[\mathscr{B},[\mathscr{A},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{B},[\mathscr{B},\mathscr{C}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{A},[\mathscr{C},\mathscr{G}]]]=\mathbf{0}, \tag{4.2.25}$$

$$[\mathscr{A},[\mathscr{B},[\mathscr{A},\mathscr{G}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{B},[\mathscr{B},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{A},[\mathscr{E},\mathscr{E}]]]=\mathbf{0}, \tag{4.2.26}$$

$$[\mathscr{A},[\mathscr{B},[\mathscr{C},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{B},[\mathscr{B},\mathscr{G}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{C},[\mathscr{B},\mathscr{C}]]]=\mathbf{0}, \tag{4.2.27}$$

$$[\mathscr{A},[\mathscr{B},[\mathscr{E},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{D},[\mathscr{A},\mathscr{B}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{C},[\mathscr{C},\mathscr{D}]]]=\mathbf{0}, \tag{4.2.28}$$

$$[\mathscr{A},[\mathscr{C},[\mathscr{A},\mathscr{B}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{D},[\mathscr{A},\mathscr{D}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{C},[\mathscr{C},\mathscr{G}]]]=\mathbf{0}, \tag{4.2.29}$$

$$[\mathscr{A},[\mathscr{C},[\mathscr{A},\mathscr{D}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{D},[\mathscr{B},\mathscr{C}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{E},[\mathscr{A},\mathscr{C}]]]=\mathbf{0}, \tag{4.2.30}$$

$$[\mathscr{A},[\mathscr{C},[\mathscr{A},\mathscr{G}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{D},[\mathscr{B},\mathscr{D}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{E},[\mathscr{A},\mathscr{E}]]]=\mathbf{0}, \tag{4.2.31}$$

$$[\mathscr{A},[\mathscr{C},[\mathscr{B},\mathscr{D}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{D},[\mathscr{B},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{E},[\mathscr{B},\mathscr{C}]]]=\mathbf{0}, \tag{4.2.32}$$

$$[\mathscr{B},[\mathscr{A},[\mathscr{A},\mathscr{B}]]]=\mathbf{0}, \quad [\mathscr{B},[\mathscr{D},[\mathscr{B},\mathscr{G}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{E},[\mathscr{C},\mathscr{D}]]]=\mathbf{0}, \tag{4.2.33}$$

$$[\mathscr{B},[\mathscr{A},[\mathscr{A},\mathscr{D}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{A},[\mathscr{A},\mathscr{C}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{E},[\mathscr{C},\mathscr{E}]]]=\mathbf{0}, \tag{4.2.34}$$

$$[\mathscr{B},[\mathscr{A},[\mathscr{B},\mathscr{D}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{A},[\mathscr{A},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{E},[\mathscr{C},\mathscr{G}]]]=\mathbf{0}, \tag{4.2.35}$$

$$[\mathscr{B},[\mathscr{A},[\mathscr{B},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{A},[\mathscr{C},\mathscr{D}]]]=\mathbf{0}, \quad [\mathscr{C},[\mathscr{E},[\mathscr{E},\mathscr{E}]]]=\mathbf{0}, \tag{4.2.36}$$

$$[\mathscr{D}_{\lambda,j},[\mathscr{B},[\mathscr{C},\mathscr{D}_{\lambda,j}]]]=\mathbf{0}, \quad [\mathscr{D}_{\lambda,j},[\mathscr{B},[\mathscr{D}_{\lambda,j},\mathscr{G}]]]=\mathbf{0}, \quad [\mathscr{E}_{\lambda,j},[\mathscr{C},[\mathscr{D},\mathscr{E}_{\lambda,j}]]]=\mathbf{0}, \tag{4.2.37}$$

$$[\mathscr{D}_{\lambda,j},[\mathscr{B},[\mathscr{D}_{\lambda,j},\mathscr{E}]]]=\mathbf{0}, \quad [\mathscr{E}_{\lambda,j},[\mathscr{C},[\mathscr{B},\mathscr{E}_{\lambda,j}]]]=\mathbf{0}, \quad [\mathscr{E}_{\lambda,j},[\mathscr{C},[\mathscr{E}_{\lambda,j},\mathscr{G}]]]=\mathbf{0}, \tag{4.2.38}$$

$$[[\mathscr{A},\mathscr{B}],[\mathscr{A},\mathscr{B}]]=\mathbf{0}, \qquad [[\mathscr{A},\mathscr{C}],[\mathscr{E},\mathscr{E}]]=\mathbf{0}, \qquad [[\mathscr{B},\mathscr{C}],[\mathscr{B},\mathscr{C}]]=\mathbf{0}, \tag{4.2.39}$$

$$[[\mathscr{A},\mathscr{B}],[\mathscr{A},\mathscr{D}]]=\mathbf{0}, \qquad [[\mathscr{A},\mathscr{D}],[\mathscr{A},\mathscr{D}]]=\mathbf{0}, \qquad [[\mathscr{B},\mathscr{D}],[\mathscr{B},\mathscr{D}]]=\mathbf{0}, \tag{4.2.40}$$

$$[[\mathscr{A},\mathscr{B}],[\mathscr{B},\mathscr{D}]]=\mathbf{0}, \qquad [[\mathscr{A},\mathscr{D}],[\mathscr{B},\mathscr{D}]]=\mathbf{0}, \qquad [[\mathscr{C},\mathscr{E}],[\mathscr{C},\mathscr{E}]]=\mathbf{0}, \tag{4.2.41}$$

$$[[\mathscr{A},\mathscr{C}],[\mathscr{A},\mathscr{C}]]=\mathbf{0}, \qquad [[\mathscr{A},\mathscr{E}],[\mathscr{A},\mathscr{E}]]=\mathbf{0}, \qquad [[\mathscr{C},\mathscr{E}],[\mathscr{E},\mathscr{E}]]=\mathbf{0}, \tag{4.2.42}$$

$$[[\mathscr{A},\mathscr{C}],[\mathscr{A},\mathscr{E}]]=\mathbf{0}, \qquad [[\mathscr{A},\mathscr{E}],[\mathscr{C},\mathscr{E}]]=\mathbf{0}, \qquad [[\mathscr{E},\mathscr{E}],[\mathscr{E},\mathscr{E}]]=\mathbf{0}, \tag{4.2.43}$$

$$[[\mathscr{A},\mathscr{C}],[\mathscr{C},\mathscr{E}]]=\mathbf{0}, \qquad [[\mathscr{A},\mathscr{E}],[\mathscr{E},\mathscr{E}]]=\mathbf{0}, \tag{4.2.44}$$

$$[[\mathscr{A},\mathscr{G}_{\lambda,j}],[\mathscr{A},\mathscr{G}_{\lambda,j}]]=\mathbf{0}, \tag{4.2.45}$$

$$[[\mathscr{B},\mathscr{E}_{\lambda,j}],[\mathscr{B},\mathscr{E}_{\lambda,j}]]=\mathbf{0}, \quad [[\mathscr{B},\mathscr{G}_{\lambda,j}],[\mathscr{B},\mathscr{G}_{\lambda,j}]]=\mathbf{0}, \quad [[\mathscr{C},\mathscr{D}_{\lambda,j}],[\mathscr{C},\mathscr{D}_{\lambda,j}]]=\mathbf{0}, \tag{4.2.46}$$

$$[[\mathscr{C},\mathscr{G}_{\lambda,j}],[\mathscr{C},\mathscr{G}_{\lambda,j}]]=\mathbf{0}, \tag{4.2.47}$$

$$[\mathscr{A},[\mathscr{F},\mathscr{F}]]=\mathbf{0}, \tag{4.2.48}$$

$$[\mathscr{A},[\mathscr{F},[\mathscr{A},\mathscr{F}]]]=\mathbf{0}, [\mathscr{F}_{\lambda,j},[\mathscr{A},[\mathscr{A},\mathscr{F}_{\lambda,j}]]]=\mathbf{0}, \qquad [[\mathscr{F},\mathscr{F}],[\mathscr{F},\mathscr{F}]]=\mathbf{0}, \tag{4.2.49}$$

$$[\mathscr{A},[\mathscr{F},[\mathscr{A},\mathscr{G}]]]=\mathbf{0}, [\mathscr{F}_{\lambda,j},[\mathscr{A},[\mathscr{F}_{\lambda,j},\mathscr{G}]]]=\mathbf{0}, [[\mathscr{A},\mathscr{F}_{\lambda,j}],[\mathscr{A},\mathscr{F}_{\lambda,j}]]=\mathbf{0}, \tag{4.2.50}$$

where each relation in (4.2.18)–(4.2.50) should be understood as a family of equalities, given the slight abuse of notation as defined just above the formulas. In particular, for each fixed numerical mesh, their appropriateness varies according to where λ lies. In detail:

- (4.2.18)–(4.2.47) play a role in (3.3.1),
- (4.2.23) play a role in (3.3.2),
- (4.2.18), (4.2.45), (4.2.48)–(4.2.50) play a part in (3.3.3).

The idea is then to remove such terms from a Hall basis.

However, since the constraints (4.2.18)–(4.2.50) are not symmetric in (4.2.2)–(4.2.8) and the terms in a Hall basis (with more than one commutator) are not symmetric in the alphabets (4.2.12)–(4.2.14), it may be possible, at least in principle, that certain bijections between (4.2.2)–(4.2.8) and (4.2.12)–(4.2.14) yield less non-zero terms than others when considering (4.2.18)–(4.2.50) on top of a Hall basis.

To put it another way, since (4.2.18)–(4.2.50) do not distinguish between different $\mathscr{A}$'s, different $\mathscr{B}$'s or different $\mathscr{C}$'s, the question can be raised equivalently as to whether specific distinct bijections between, respectively, (4.2.12), (4.2.13), (4.2.14) and

$$\left\{\mathscr{A}^{\boldsymbol{f}},\mathscr{A}^{\boldsymbol{f}},\mathscr{A}^{\boldsymbol{f}},\mathscr{A}^{\boldsymbol{f}},\mathscr{B}^{\boldsymbol{f}},\mathscr{B}^{\boldsymbol{f}},\mathscr{B}^{\boldsymbol{f}},\mathscr{B}^{\boldsymbol{f}},\mathscr{C}^{\boldsymbol{f}},\mathscr{C}^{\boldsymbol{f}},\mathscr{D}^{\boldsymbol{f}},\mathscr{E}^{\boldsymbol{f}}_{\lambda,1},\mathscr{E}^{\boldsymbol{f}}_{\lambda,2},\mathscr{E}^{\boldsymbol{f}}_{\lambda,3},\mathscr{G}^{\boldsymbol{f}}\right\}, \tag{4.2.51}$$

$$\left\{\mathscr{B}^{\boldsymbol{\iota}},\mathscr{C}^{\boldsymbol{\iota}},\mathscr{D}^{\boldsymbol{\iota}},\mathscr{E}^{\boldsymbol{\iota}},\mathscr{G}^{\boldsymbol{\iota}}\right\}, \tag{4.2.52}$$

$$\left\{\mathscr{A}^{\boldsymbol{g}},\mathscr{A}^{\boldsymbol{g}},\mathscr{A}^{\boldsymbol{g}},\mathscr{A}^{\boldsymbol{g}},\mathscr{A}^{\boldsymbol{g}},\mathscr{A}^{\boldsymbol{g}},\mathscr{F}^{\boldsymbol{g}}_{\lambda,1},\mathscr{F}^{\boldsymbol{g}}_{\lambda,2},\mathscr{F}^{\boldsymbol{g}}_{\lambda,3},\mathscr{G}^{\boldsymbol{g}}_{\lambda,1},\mathscr{G}^{\boldsymbol{g}}_{\lambda,2},\mathscr{G}^{\boldsymbol{g}}_{\lambda,3}\right\}, \tag{4.2.53}$$

yield less non-zero terms than others when considering (4.2.18)–(4.2.50) on top of a Hall basis.

This is indeed the case, and having this in mind, when faced with the task of decreasing the volume of linear algebra in the exact integration of (4.2.1), one should then search for a bijection from, respectively, (4.2.12)–(4.2.14) to (4.2.51)–(4.2.53), that minimizes the number of non-zero terms in a Hall basis when taking (4.2.18)–(4.2.50) into account.

The last equivalence is then particularly useful given that the number of distinct bijections between elements with repetition is often much smaller than the total number of bijections.

If even the number of distinct permutation is so large that it is not practical to search for a minimizer by brute-force, one might benefit from simulated annealing (Press, Teukolsky, Vetterling and Flannery, 2007, Section 10.12).

As recorded in Table 4.1 under "Reduced Hall basis with non-zero integrands", the second mechanism leads to substantial savings when compared with a vanilla Hall basis that does not take (4.2.18)–(4.2.50) into account.

Finally, the third mechanism above follows simply from the fact that, as stated, certain linear combinations between various integrands integrate exactly to zero. Thus, one should

search such occurrences to further decrease the volume of linear algebra in the exact integration of (4.2.1), as depicted in Table 4.1 under "Reduced Hall basis with non-zero integrals".

To conclude, it is intriguing to observe in Table 4.1 that, when compared with the rest of the eigenvalue range (3.3.2)–(3.3.3), it is the intermediary regime (3.3.1), which, as discussed in Chapter 1, is not covered by alternative techniques, that requires the largest volume of linear algebra per evaluation of $\lambda \mapsto \tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,j,\mathcal{T}_{l-1}}(c_k, c_{k+1})$.

4.2.2 Self-adjoint basis and graded FLA

In passing, it is important to say at this point that, in principle, one could call upon yet another mechanism to further reduce the number of commutators and volume of linear algebra. This is the theory of the graded FLA introduced by the self-adjoint basis put forth in (Munthe–Kaas and Owren, 1999, Subsection 4.a) and further discussed in (Iserles, Munthe–Kaas, Nørsett and Zanna, 2000, Subsection 5.2), for settings without oscillatory behaviour.

While preparing this dissertation, these ideas were implemented and tested on Fer streamers, in the eigenvalue range (3.3.3). However, it has been found that the graded FLA induced by the self-adjoint basis in essence destroys the advantages gained from the deliberate interpolation of $\boldsymbol{B}^{\text{fine}}_{\lambda,1}(c_k, c_k + h_k t)$ in Definition 4.1.1 at the right-boundary point, $t = 1$, that, as explained carefully in the previous Chapter 3, reduces the quadrature error in many cases when high oscillation is present. In addition, by construction, the self-adjoint basis and its graded FLA in (Munthe–Kaas and Owren, 1999; Iserles, Munthe–Kaas, Nørsett and Zanna, 2000) are advantageous when the step size h is close to 0, but, in practice, it often happens that Fer streamers work well even with h near to 1, so that one cannot reap the benefits from the self-adjoint basis, unless needlessly reducing h, and thereby increasing the number of function evaluations of q without need, which is not desirable since these can be of considerable cost in practice. Because of all this, although implemented, this graded FLA was discarded from the Fer streamers' MATLAB package that accompanies the following Chapter 5.

4.3 Conclusions

Having reviewed the approximation properties, carefully developed in Chapters 2 and 3, for the truncated and discretized flow $\tilde{\tilde{\boldsymbol{F}}}_{\lambda,n}(c_k, c_{k+1})$ and solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, which, by construction, can be computed exactly, we have seen in this chapter that their implementation in practice, benefits from three mechanisms that lead to a reduced Hall basis. As demonstrated clearly in Table 4.1, this is a very useful concept, since it yields a significant decrease in the amount of linear algebra required for their practical implementation, with uniform global orders 4, 7, 10 and 13, with respect to (1.1.6)–(1.1.7), which has now been

			Eigenvalue range (3.3.1)				Eigenvalue range (3.3.2)				Eigenvalue range (3.3.3)			
			Free magma	Hall basis	Reduced Hall basis with non-zero integrands (4.2.18) up to (4.2.47)	Reduced Hall basis with non-zero integrals	Free magma	Hall basis	Reduced Hall basis with non-zero integrands (4.2.23)	Reduced Hall basis with non-zero integrals	Free magma	Hall basis	Reduced Hall basis with non-zero integrands (4.2.18), (4.2.45), (4.2.48) up to (4.2.50)	Reduced Hall basis with non-zero integrals
global order	4	$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_3}$	15	15	15	15	5	5	5	5	12	12	12	12
	7	$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_6}$	27	27	27	27	8	8	8	8	24	24	24	24
		$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,2,\mathcal{T}_3}$	225	105	92	92	25	10	10	10	144	66	51	49
	10	$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_9}$	39	39	39	39	11	11	11	11	36	36	36	36
		$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,2,\mathcal{T}_6}$	729	351	308	308	64	28	28	28	576	276	210	205
		$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,3,\mathcal{T}_3}$	3375	1120	≤ 412	≤ 412	125	40	38	34	1728	572	≤ 408	≤ 404
	13	$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,1,\mathcal{T}_{12}}$	51	51	51	51	14	14	14	14	48	48	48	48
		$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,2,\mathcal{T}_9}$	1521	741	650	650	121	55	55	55	1296	630	477	469
		$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,3,\mathcal{T}_6}$	19683	6552	≤ 5072	≤ 5072	512	168	166	154	13824	4600	≤ 3236	≤ 3201
		$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,4,\mathcal{T}_3}$	50625	12600	≤ 7899	≤ 7867	625	150	134	126	20736	5148	≤ 3069	≤ 3060
		$\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,5,\mathcal{T}_3}$	50625	5460	≤ 3444	≤ 3438	625	45	45	42	20736	2145	≤ 1182	≤ 1082

Table 4.1: Number of terms in each integrand of $\tilde{\boldsymbol{I}}^{\text{fine}}_{\lambda,j,\mathcal{T}_{l-1}}$, for each eigenvalue range and different global orders

realized in the form of a MATLAB package, which is presented, with several illustrative examples, in the following Chapter 5.

Chapter 5

Fer streamers' MATLAB package

The present chapter discusses some of the implementation details of the Fer streamers' MATLAB package that accompanies this dissertation, which is based on the work in the previous Chapters 2, 3 and 4, and can be found in (Ramos, 2015*c*). In particular, the package embodies the approximations of the truncated and discretized solution $\tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(c_{k+1})$, with global orders $4, 7, 10$ and 13, uniform over the entire eigenvalue spectrum, discussed throughout the previous chapters. As mentioned earlier in Chapter 1, these uniform and high-order approximations can be employed together with the two different eigenvalue representations given via $\lambda \mapsto \eta_\lambda$ in Theorem 1.0.1 and $\lambda \mapsto \theta_\lambda(b)$ in Theorem 1.0.2, along with root-finding techniques, to approximate the eigenvalues of regular Sturm–Liouville problems (1.0.1)–(1.0.2), with continuous and piecewise analytic potentials (1.0.13).

The implementation aspects in the current chapter concern the specific choices of eigenvalue representations and root-finding tools used in the MATLAB package, along with heuristics for mesh selection and error estimation. For ease of use, a description of how to call the package is also provided, which is then illustrated with several numerical results.

5.1 Eigenvalue characterizations via Prüfer's scaled variables

In order to approximate the eigenvalues via value or index, i.e., by (a) or (b) in page 1, the MATLAB package that comes with this thesis is based on the representation with $\lambda \mapsto \theta_\lambda(b)$ in Theorem 1.0.2[1], rather than on the one with $\lambda \mapsto \eta_\lambda$ in Theorem 1.0.1.

This is done because, as explained already in Chapter 1, $\lambda \mapsto \theta_\lambda(b)$ is strictly increasing and provides the j-th eigenvalue as its pre-image of $\beta + j\pi$. On the contrary, $\lambda \mapsto \eta_\lambda$ is oscillatory with roots equal to the eigenvalues, which does not give information about the indices of the eigenvalues and therefore cannot be used to solve problem (b) in page 1.

[1] To be precise, the Fer streamers' MATLAB package employs more elaborate versions of $\lambda \mapsto \theta_\lambda(b)$ and Theorem 1.0.2, which use scaled rather than unscaled Prüfer variables (Pryce, 1993, Section 5; Zettl, 2005, p. 81–87). More concretely, it uses modified versions of the stabilized algorithm in (Pruess and Fulton, 1993, p. 364–367) and of the Prüfer transformation in (Ixaru, De Meyer and Berghe, 1999, p. 263–265).

As mentioned also in Chapter 1, the uniform approximations $(\lambda, t) \mapsto \tilde{\tilde{\boldsymbol{Y}}}_{\lambda,n}(t)$ to (1.0.9) developed in Chapters 2–4, can be used to approximate $\lambda \mapsto \theta_\lambda(b)$, independently of λ (Pruess and Fulton, 1993, p. 364–367; Ixaru, De Meyer and Berghe, 1999, p. 263–265).

It is important to note that like the exact $\lambda \mapsto \theta_\lambda(b)$, the approximation $\lambda \mapsto \tilde{\tilde{\theta}}_{\lambda,n}(b)$ is also strictly increasing.

With an approximation $\lambda \mapsto \tilde{\tilde{\theta}}_{\lambda,n}(b)$ in hand, one can then approximate λ_j, the solution to (1.0.8), by $\tilde{\tilde{\lambda}}_{j,n}$, the solution to

$$\tilde{\tilde{\theta}}_{\lambda,n}(b) = \beta + j\pi. \tag{5.1.1}$$

Since, in general, one cannot solve (5.1.1) exactly, i.e., root-find $\lambda \mapsto (\tilde{\tilde{\theta}}_{\lambda,n}(b) - \beta - j\pi)$ in closed-form, instead, as discussed in the next subsection, one calls upon a root-finding algorithm, which approximates $\tilde{\tilde{\lambda}}_{j,n}$ by $\tilde{\tilde{\lambda}}_{j,n}$, up to prescribed tolerance.

5.2 Root-finding via Brent's method

As mentioned in the previous subsection, the Fer streamers' MATLAB package outputs an approximation $\tilde{\tilde{\lambda}}_{j,n}$ to $\tilde{\tilde{\lambda}}_{j,n}$, which is the result of applying a root-finding algorithm to $\lambda \mapsto (\tilde{\tilde{\theta}}_{\lambda,n}(b) - \beta - j\pi)$, up to stipulated tolerance.

Given that $\lambda \mapsto \tilde{\tilde{\theta}}_{\lambda,n}(b)$ is strictly increasing as explained in the previous subsection, to achieve prescribed tolerance in this specific root-find one can proceed with standard bisection: first bracket the unique root in a initial interval, then halve the interval with bisection, an take the subinterval which is guaranteed to contain the unique root; repeat this procedure until the length of the current interval is smaller than the requested tolerance.

Bisection with an increasing function works well because it divides the interval while making sure that the chosen subinterval contains the root. Unfortunately, it is rather slow.

Hence, instead the Fer streamers' MATLAB package employs a faster version of bisection, known as Brent's method (Brent, 2002, Section 4), which shares all the properties of the bisection, but is much faster.

5.3 Heuristics for mesh selection and error estimation

The Fer streamers' MATLAB package uses a nested rule with uniform global orders $\{g_1, g_2, g_3\}$, where $g_1 < g_2 < g_3$. There are two options in the current version: $\{g_1, g_2, g_3\} := \{4, 7, 10\}$ and $\{g_1, g_2, g_3\} := \{7, 10, 13\}$; the latter being the default. Since the interpolation points in Definition 4.1.1 are nested, i.e., $\mathcal{T}_3 \subseteq \mathcal{T}_6 \subseteq \mathcal{T}_9 \subseteq \mathcal{T}_{12}$, nested rules do not incur extra function evaluations of the potential q.

5.3.1 Mesh selection

The motivation for the mesh selection is straightforward: keep to a minimum the number of function evaluations of the potential q. In particular, the Fer streamers' MATLAB package employs a modified version of the mesh selection in (Ixaru, De Meyer and Berghe, 1997, p. 305–306) and (Ledoux, Daele and Berghe, 2010, p. 764–765).

In short, for each $[c_k, c_{k+1}^{\text{trial}}]$, the mesh selection is based on a local difference between Fer streamers with uniform local orders $\{g_2 + 1, g_3 + 1\}$, which is tested on

$$\lambda \in \left\{ q_{\min}, \frac{\int_{c_k}^{c_{k+1}^{\text{trial}}} q(\xi) d\xi}{c_{k+1}^{\text{trial}} - c_k}, q_{\max} \right\}, \tag{5.3.1}$$

where $q_{\min}$ only needs to be a lower bound for the minimum of the potential and $q_{\max}$ only needs to be an upper bound for its maximum. With the uniform guarantees from Theorem 4.1.1, the specific choice (5.3.1) is motivated by the magnitude and behaviour of the first Fer streamer in Remark 2.1.5, discussed at length in Section 4.1. Once computed, the numerical mesh remains unaltered from start to finish.

5.3.2 Error estimation

Having computed the mesh, the Fer streamers' MATLAB package first brackets the eigenvalues with uniform global order g_1. With these preliminary brackets, the package then runs with uniform global orders g_2 and g_3. The error estimation of the absolute error and relative error are then given by, respectively, the absolute and relative errors between

$$\tilde{\tilde{\lambda}}_{j,(\log((g_2+2)/3)/\log(2))} \text{ and } \tilde{\tilde{\lambda}}_{j,(\log((g_3+2)/3)/\log(2))}. \tag{5.3.2}$$

5.4 Calling the Fer streamers MATLAB package

The Fer streamers' MATLAB package can be downloaded from (Ramos, 2015*c*). The root file `m_index.m` sets up the input, calls the main file `m_Fer_streamers.m` and provides the output.

5.4.1 Input

To run the main file `m_Fer_streamers.m`, the root file `m_index.m` sets up the input:

- `parameter`:

which serves to parameterize the Sturm–Liouville problem if necessary, otherwise set to empty,

- a, b, q, $q_{\min}$, $q_{\max}$, α_1, α_2, β_1, β_2:

which characterize the Sturm–Liouville problem, where $q_{\min}$ only needs to be a lower bound for the minimum of the potential and $q_{\max}$ only needs to be an upper bound for its maximum,

- `index_range_to_eigenvalues_or_eigenvalue_range_to_eigenvalues`, `range_min`, `range_max`:

that set up whether the eigenvalues should be computed according to their indices or values, and on which range, as well as,

- `error_absolute_or_relative`, `tol_stopping_criteria`:

that set up the tolerance and type of error that should be used for the stopping criteria.

5.4.2 Output

The main file `m_Fer_streamers.m` outputs:

- `all_t_and_q_at_t_pairs`:

a two column matrix that contains per line all evaluations $(t, q(t))$ used or discarded from start to finish,

- `eigenvalues_indices_absoluteErrors_relativeErrors`:

a four column matrix which collects per line the requested eigenvalues, their indices and an error estimation of the absolute and relative errors in their approximation, as described in Subsection 5.3.2.

5.5 Numerical results

To illustrate the Fer streamers' MATLAB package with nested uniform global orders $\{7, 10, 13\}$ (c.f., Section 5.3), the numerical results in this section describe its output on the four Sturm–Liouville problems (3.5.1)–(3.5.4) below, when set to approximate their first 500 eigenvalues:

- `index_range_to_eigenvalues_or_eigenvalue_range_to_eigenvalues` =`'index_range_to_eigenvalues'`, `range_min`=0, `range_max`=499,

up to prescribed absolute error with tolerance 10^{-8}:

- `error_absolute_or_relative`=`'absolute'`, `tol_stopping_criteria`=10^{-8}.

To cover different phenomena, this sections then examines:

- the Anderssen and de Hoog problem (3.5.1),

- the second Paine problem (3.5.2),
- the Coffey–Evans problem (3.5.3),
- the truncated Gelfand–Levitan problem (3.5.4).

The numerical results in Figures 5.1–5.2 illustrate the output from Fer streamers' MATLAB package with nested uniform global orders $\{7, 10, 13\}$ and absolute error tolerance 10^{-8} for the first 500 eigenvalues of the Sturm–Liouville problems (3.5.1), (3.5.2), (3.5.3) and (3.5.4). Apart from the number of function evaluations of the potential q used or discarded by Fer streamers, each plot displays the estimated absolute/relative error by Fer streamers as defined in Subsection 5.3.2 together with the actual absolute/relative error when compared with a reference solution, computed with MATSLISE's package (Ledoux, Daele and Berghe, 2005). Absent circles mean that the estimated error by Fer streamers is equal to zero, i.e., that (5.3.2) with $g_2 = 10$ and $g_3 = 13$ are identical in machine precision. In particular, one can see that the estimated error by Fer streamers is in-line with the prescribed absolute error tolerance 10^{-8}, being often quite conservative.

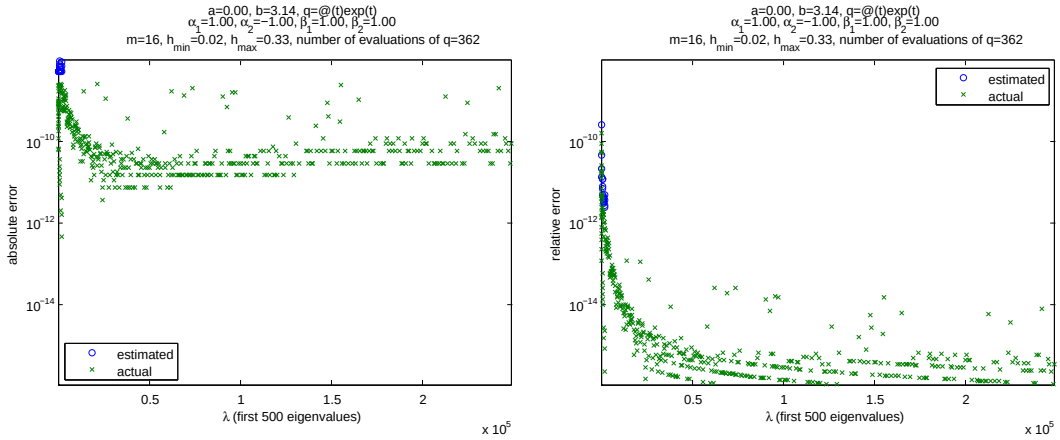


Figure 5.1: Absolute error (left) and relative error (right) with Fer streamers' MATLAB package with nested uniform global orders $\{7, 10, 13\}$ and absolute error tolerance 10^{-8} for the first 500 eigenvalues of the Anderssen and de Hoog problem (3.5.1). Apart from the number of evaluations of q used or discarded by Fer streamers, each plot displays the estimated error by Fer streamers as defined in Subsection 5.3.2 together with the actual error when compared with a reference solution. Absent circles signify that the estimated error by Fer streamers equals zero, i.e., that (5.3.2) with $g_2 = 10$ and $g_3 = 13$ coincide up to machine precision.

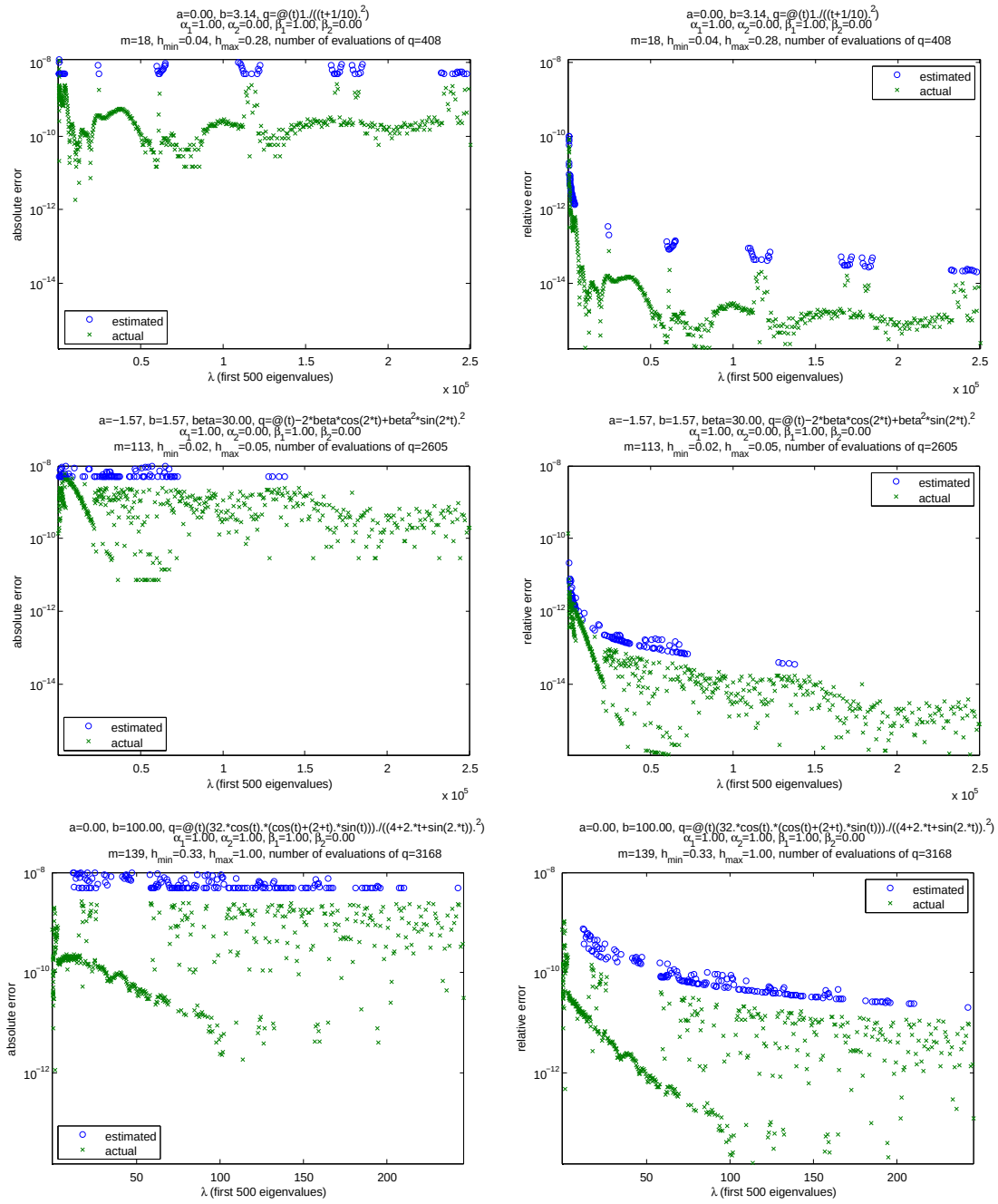


Figure 5.2: Absolute error (left) and relative error (right) with Fer streamers' MATLAB package with nested uniform global orders $\{7, 10, 13\}$ and absolute error tolerance 10^{-8} for the first 500 eigenvalues of the second Paine problem (3.5.2) (top), the Coffey–Evans problem (3.5.3) (middle), the truncated Gelfand–Levitan problem (3.5.4) (bottom). Apart from the number of evaluations of q used or discarded by Fer streamers, each plot displays the estimated error by Fer streamers as defined in Subsection 5.3.2 together with the actual error when compared with a reference solution. Absent circles signify that the estimated error by Fer streamers equals zero, i.e., that (5.3.2) with $g_2 = 10$ and $g_3 = 13$ coincide up to machine precision.

5.6 Conclusions

We have discussed in this chapter several implementation details that surround the Fer streamers' MATLAB package, that accompanies this dissertation, and places the theoretical work in Chapters 2, 3 and 4 into practice.

We have seen throughout the numerical results presented within, that the output which approximates the eigenvalues of regular Sturm–Liouville problems (1.0.1)–(1.0.2), with continuous and piecewise analytic potentials (1.0.13), provided by the package, is in line with the input tolerance, being often quite conservative.

In addition, we have also seen that Fer streamers perform well with large step sizes and small number of evaluations of the potential function.

The final output of Chapters 2–5 is then an algorithm that approximates the eigenvalues of regular Sturm–Liouville problems based on Fer streamers, which is mathematically guaranteed to be uniformly precise and most affordable, throughout all orders of magnitude of the eigenvalues.

Chapter 6

A generalized truncation

Motivated by the discussion in Section 1.2, the present chapter extends the basic results that support the work in Chapter 2, from the classical setting with continuous and piecewise analytic potentials (1.0.13) to the general case with absolutely integrable potentials (1.0.14), as reported in (Ramos, 2014).

Namely, we generalize Theorems 2.1.4–2.1.5, from (1.0.13) to (1.0.14), and prove that they retain the same advantageous properties either in the original setting or in the general case, since they remain accompanied by error bounds which have the properties (i) and (ii) from page 25.

In particular, in the present chapter, we establish precisely the manner in which the regularity of the potential influences Theorems 2.1.4–2.1.5 through:

- the maximum step size, and,
- the converge rate.

In Section 6.1, we revisit and extend Assumption 1.1.1. In particular, given the immense variety of the regularity of the potentials across the general set (1.0.14), we identify four classes of regularity that cover the entire set (1.0.14), but restrict the maximum step size differently in view of the specific characteristics of the potentials across the four classes.

Accordingly, to account for each of the four classes, in Section 6.2, we extend the methodology in Chapter 2 that partitions the eigenvalue range (2.0.1), as we put forth a generalization of the two uniform regimes (1.1.6)–(1.1.7).

Finally, in Section 6.3, we explain how the regularity of the potential influences, firstly: the magnitude of the Fer streamers from Theorem 2.1.3, in the extended version of Theorem 2.1.4, and secondly: the convergence rate of the local and global truncation errors from Definition 2.1.6, in the generalized version of Theorem 2.1.5, that opens the door to a future extension of the numerical method put forth in this dissertation, to broader settings.

6.1 Four classes of potentials

Towards generalizing Theorems 2.1.4–2.1.5, it is useful to cluster the set of $\mathrm{L}^1([a,b],\mathbb{R})$ potentials in four nested classes according to their regularity. In particular, it is of the utmost importance to identify the largest

$$p \in [1,\infty]$$

such that

$$q \in \mathrm{L}^p([a,b],\mathbb{R}).$$

Class I (Essentially Piecewise Absolutely Continuous Potentials)**.** *A potential q is said to belong to this class if*

$$p = \infty$$

and there exist

$$m \in \mathbb{Z}^+, \tag{6.1.1}$$

$$c_0 := a < c_1 < \cdots < c_{m-1} < c_m := b, \tag{6.1.2}$$

$$h_{\min} := \min_{k\in\{0,1,\ldots,m-1\}} \{c_{k+1} - c_k\}, \tag{6.1.3}$$

$$h_{\max} := \max_{k\in\{0,1,\ldots,m-1\}} \{c_{k+1} - c_k\}, \tag{6.1.4}$$

$$\gamma \in [1,\infty], \tag{6.1.5}$$

$$q_0 \in \mathrm{AC}([c_0,c_1],\mathbb{R}), \ldots, q_{m-1} \in \mathrm{AC}([c_{m-1},c_m],\mathbb{R}), \tag{6.1.6}$$

such that, for all $k \in \{0,1,\ldots,m-1\}$,

$$q_k' \in \mathrm{L}^\gamma([c_k,c_{k+1}],\mathbb{R}), \tag{6.1.7}$$

$$q(t) = q_k(t) \ \text{ a.e. } \ t \in [c_k,c_{k+1}]. \tag{6.1.8}$$

In this case, it is assumed that the numerical mesh (6.1.1)–(6.1.4) has been refined in such a way that

$$\lambda \geq \mathit{ess}\inf\{q\} \implies h_{\max} \leq (\mathit{ess}\sup\{q\} - \mathit{ess}\inf\{q\})^{-\frac{1}{2}}, \tag{6.1.9}$$

$$\lambda < \mathit{ess}\inf\{q\} \implies h_{\max}^2(\mathit{ess}\sup\{q\} - \lambda) \leq 1, \tag{6.1.10}$$

$$\frac{h_{\max}}{h_{\min}} \leq 2, \tag{6.1.11}$$

and that the big $\mathcal{O}$ notation and the small o notation refers to one of the two uniform

regimes

$$h_{\max} \to 0^+, \ \textit{uniformly w.r.t.} \ |\lambda - \operatorname{ess\,sup}\{q\}| \leq h_{\max}^{-2}, \tag{6.1.12}$$

$$h_{\max} \to 0^+, \ \textit{uniformly w.r.t.} \ \ \lambda - \operatorname{ess\,sup}\{q\} \ \geq h_{\max}^{-2}. \tag{6.1.13}$$

Class II (Essentially Bounded Potentials)**.** *A potential q is said to belong to this class if*

$$p = \infty.$$

In this case, it is assumed that the numerical mesh (6.1.1)–(6.1.4) *is such that* (6.1.9), (6.1.10) *and* (6.1.11) *hold true and that the big $\mathcal{O}$ notation and the small o notation refers to one of the two uniform regimes* (6.1.12) *and* (6.1.13).

Class III. *A potential q is said to lie in this class if*

$$p \in (1, \infty).$$

In this case, it is assumed that the numerical mesh (6.1.1)–(6.1.4) *is such that*

$$\lambda \geq 0 \quad \Longrightarrow \quad h_{\max} \leq \left(4\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)^{-\frac{p}{2p-1}}, \tag{6.1.14}$$

$$\lambda < 0 \quad \Longrightarrow \quad h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})} + h_{\max}^2 |\lambda| \leq 1, \tag{6.1.15}$$

$$\frac{h_{\max}}{h_{\min}} \leq 2, \tag{6.1.16}$$

and that the big $\mathcal{O}$ notation and the small o notation refers to one of the two uniform regimes

$$h_{\max} \to 0^+, \ \textit{uniformly w.r.t.} \ |\lambda| \leq h_{\max}^{-2} \left(1 - h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right), \tag{6.1.17}$$

$$h_{\max} \to 0^+, \ \textit{uniformly w.r.t.} \ \ \lambda \ \geq h_{\max}^{-2} \left(1 - h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right). \tag{6.1.18}$$

Class IV (Absolutely Integrable Potentials)**.** *A potential q is said to lie in this class if*

$$p = 1.$$

In this case, it is assumed that the numerical mesh (6.1.1)–(6.1.4) *is such that*

$$c_1, \ldots, c_{m-1} \ \textit{are Lebesgue points of } q,$$

that (6.1.14), (6.1.15) *and* (6.1.16) *hold true with $p = 1$, and that the big $\mathcal{O}$ notation and*

the small o notation refers to one of the two uniform regimes (6.1.17) *and* (6.1.18) *with* $p = 1$.

6.2 Extended methodology

As in Chapter 2, our approach consists of a three-step procedure, which we extend in this section to account for (1.0.14) in addition to (1.0.13). Firstly, we extend (2.0.1). In particular, when dealing with potentials in Classes I or II, according to the two cases distinguished in (6.1.9) and (6.1.10), the eigenvalue interval

$$\Big[\operatorname{ess\,sup}\{q\} - h_{\max}^{-2}, +\infty\Big)$$

is divided into the two pieces

$$\lambda \in \Big[\operatorname{ess\,sup}\{q\} - h_{\max}^{-2}, \operatorname{ess\,sup}\{q\} + h_{\max}^{-2}\Big] \cup \Big[\operatorname{ess\,sup}\{q\} + h_{\max}^{-2}, +\infty\Big),$$

and when dealing with potentials in Classes III or IV, in line with the two cases distinguished in (6.1.14) and (6.1.15), the eigenvalue interval

$$\Big[- h_{\max}^{-2}\Big(1 - h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\Big), +\infty\Big)$$

is divided into the two pieces

$$\begin{aligned}\lambda \in &\Big[- h_{\max}^{-2}\Big(1 - h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\Big), h_{\max}^{-2}\Big(1 - h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\Big)\Big] \\ &\cup \Big[h_{\max}^{-2}\Big(1 - h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\Big), +\infty\Big).\end{aligned}$$

Then, we approximate the solution of (1.0.4) with initial condition (1.0.5) in the two uniform regimes (6.1.12) and (6.1.13) when dealing with potentials in Classes I or II, and in the two uniform regimes (6.1.17) and (6.1.18) when dealing with potentials in Classes III or IV.

6.3 Error estimates

We now present the main contribution of the current chapter viz. an extension of Theorems 2.1.4–2.1.5 from (1.0.13) to (1.0.14).

We begin with a definition, which encapsulates the regularity of potentials in the four classes and respective uniform regimes:

Definition 6.3.1. *Let*

$$\epsilon_1 := \begin{cases} 2h_{\max} & \Leftarrow \textit{Classes I or II, and uniform regime } (6.1.12), \\ (\lambda - \mathit{ess}\sup\{q\})^{-\frac{1}{2}} & \Leftarrow \textit{Classes I or II, and uniform regime } (6.1.13), \\ 2h_{\max} & \Leftarrow \textit{Classes III or IV, and uniform regime } (6.1.17), \\ 2\lambda^{-\frac{1}{2}} & \Leftarrow \textit{Classes III or IV, and uniform regime } (6.1.18), \end{cases}$$

and

$$\epsilon_2 := \begin{cases} \frac{3}{4}\|q'\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}h_{\max}^2 & \Leftarrow \textit{Class I and } \gamma = \infty, \\ \frac{(3\gamma-1)\gamma}{(2\gamma-1)^2}\|q'\|_{\mathrm{L}^\gamma([a,b],\mathbb{R})}o\left(h_{\max}^{\frac{2\gamma-1}{\gamma}}\right) & \Leftarrow \textit{Class I and } \gamma \in (1,\infty), \\ 2\|q'\|_{\mathrm{L}^1([a,b],\mathbb{R})}o\left(h_{\max}\right) & \Leftarrow \textit{Class I and } \gamma = 1, \\ 2\|q\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}h_{\max} & \Leftarrow \textit{Class II}, \\ \frac{2p-1}{p-1}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}o\left(h_{\max}^{\frac{p-1}{p}}\right) & \Leftarrow \textit{Class III}, \\ \|q\|_{\mathrm{L}^1([a,b],\mathbb{R})}o\left(1\right) & \Leftarrow \textit{Class IV}. \end{cases}$$

The following result generalizes Theorem 2.1.4:

Theorem 6.3.1. *If q is in Class I, II, III or IV, and $l \in \mathbb{Z}^+$, then,*

$$e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}\cdots e^{\boldsymbol{D}_{\lambda,0}(a,c_1)} = \begin{bmatrix} \mathcal{O}\left(1\right) & \mathcal{O}\left(\epsilon_1\right) \\ \mathcal{O}\left(\epsilon_1^{-1}\right) & \mathcal{O}\left(1\right) \end{bmatrix},$$

$$\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right) = \epsilon_2^{2^{l-1}}\epsilon_1^{2^{l-1}-1}\begin{bmatrix} \mathcal{O}\left(\epsilon_1\right) & \mathcal{O}\left(\epsilon_1^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

where ϵ_1 and ϵ_2 vary according to the regularity of the potential as well as to the various uniform regimes, as prescribed in Definition 6.3.1.

Proof. See the Section 6.5. □

The following result generalizes Theorem 2.1.5:

Theorem 6.3.2. *If q is in Class I, II, III or IV, and $n \in \mathbb{Z}^+$, then*

$$\boldsymbol{\pi}\left(\boldsymbol{L}_{\lambda,n}^{trun.}(c_k,c_{k+1})\right) = \epsilon_2^{2^n}\epsilon_1^{2^n-1}\begin{bmatrix} \mathcal{O}\left(\epsilon_1\right) & \mathcal{O}\left(\epsilon_1^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}_{\lambda,n}^{trun.}(a,c_{k+1})\right) = h_{\max}^{-1}\epsilon_2^{2^n}\epsilon_1^{2^n-1}\begin{bmatrix} \mathcal{O}\left(\epsilon_1\right) & \mathcal{O}\left(\epsilon_1^2\right) & \mathcal{O}\left(1\right) \end{bmatrix}^\top,$$

where ϵ_1 and ϵ_2 vary according to the regularity of the potential as well as to the various uniform regimes, as prescribed in Definition 6.3.1.

Proof. See Section 6.6. □

Theorem 6.3.2 can now be specialized to the following notable cases. These are important to write down, since they illustrate the role played by the regularity of the potential.

Corollary 6.3.1. *If q is in Class I, $\gamma = \infty$ and $n \in \mathbb{Z}^+$, then, in the uniform regimes (6.1.12)–(6.1.13),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{trun.}_{\lambda,n}(c_k, c_{k+1})\right) = \left(\frac{3}{4}\|q'\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{3\times 2^n-1}\begin{bmatrix}\mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right)\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{trun.}_{\lambda,n}(a, c_{k+1})\right) = \left(\frac{3}{4}\|q'\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{3\times 2^n-2}\begin{bmatrix}\mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right)\end{bmatrix}^\top.$$

Corollary 6.3.2. *If q is in Class I, $\gamma \in (1, \infty)$ and $n \in \mathbb{Z}^+$, then, in the uniform regimes (6.1.12)–(6.1.13),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{trun.}_{\lambda,n}(c_k, c_{k+1})\right) = \left(\frac{(3\gamma-1)\gamma}{(2\gamma-1)^2}\|q'\|_{\mathrm{L}^\gamma([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{\frac{3\gamma-1}{\gamma}\times 2^n-1}\begin{bmatrix}o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{trun.}_{\lambda,n}(a, c_{k+1})\right) = \left(\frac{(3\gamma-1)\gamma}{(2\gamma-1)^2}\|q'\|_{\mathrm{L}^\gamma([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{\frac{3\gamma-1}{\gamma}\times 2^n-2}\begin{bmatrix}o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top.$$

Corollary 6.3.3. *If q is in Class I, $\gamma = 1$ and $n \in \mathbb{Z}^+$, then, in the uniform regimes (6.1.12)–(6.1.13),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{trun.}_{\lambda,n}(c_k, c_{k+1})\right) = \left(2\|q'\|_{\mathrm{L}^1([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{2\times 2^n-1}\begin{bmatrix}o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{trun.}_{\lambda,n}(a, c_{k+1})\right) = \left(2\|q'\|_{\mathrm{L}^1([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{2\times 2^n-2}\begin{bmatrix}o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top.$$

Corollary 6.3.4. *If q is in Class II and $n \in \mathbb{Z}^+$, then, in the uniform regimes (6.1.12)–(6.1.13),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{trun.}_{\lambda,n}(c_k, c_{k+1})\right) = \left(2\|q\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{2\times 2^n-1}\begin{bmatrix}\mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right)\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{trun.}_{\lambda,n}(a, c_{k+1})\right) = \left(2\|q\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{2\times 2^n-2}\begin{bmatrix}\mathcal{O}\left(h_{\max}\right) & \mathcal{O}\left(h_{\max}^2\right) & \mathcal{O}\left(1\right)\end{bmatrix}^\top.$$

Corollary 6.3.5. *If q belongs to Class III and $n \in \mathbb{Z}^+$, then, in the uniform regimes (6.1.17)–(6.1.18),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{trun.}_{\lambda,n}(c_k, c_{k+1})\right) = \left(\frac{2p-1}{p-1}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{\frac{2p-1}{p}\times 2^n - 1}\begin{bmatrix} o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{trun.}_{\lambda,n}(a, c_{k+1})\right) = \left(\frac{2p-1}{p-1}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{\frac{2p-1}{p}\times 2^n - 2}\begin{bmatrix} o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top.$$

Corollary 6.3.6. *If q belongs to Class IV and $n \in \mathbb{Z}^+$, then, in the uniform regimes (6.1.17)–(6.1.18),*

$$\boldsymbol{\pi}\left(\boldsymbol{L}^{trun.}_{\lambda,n}(c_k, c_{k+1})\right) = \left(\|q\|_{\mathrm{L}^1([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{1\times 2^n - 1}\begin{bmatrix} o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top,$$

$$\boldsymbol{\pi}\left(\boldsymbol{G}^{trun.}_{\lambda,n}(a, c_{k+1})\right) = \left(\|q\|_{\mathrm{L}^1([a,b],\mathbb{R})}\right)^{2^n} h_{\max}^{1\times 2^n - 2}\begin{bmatrix} o\left(h_{\max}\right) & o\left(h_{\max}^2\right) & o\left(1\right)\end{bmatrix}^\top.$$

Two observations are in order at this point. Firstly, one should note that Corollary 6.3.1 recovers the error bounds from the original Theorem 2.1.5, as written in Corollary 2.1.1. Secondly, one should remark that as the regularity of the potential decreases throughout Corollaries 6.3.1, 6.3.2, 6.3.3, 6.3.4, 6.3.5 and 6.3.6, so does the rate of convergence of the local and global truncation errors, starting from Corollary 6.3.1 with

$$h_{\max}^{3\times 2^n - 1} \quad \text{and} \quad h_{\max}^{3\times 2^n - 2},$$

and decreasing to Corollary 6.3.6 with

$$h_{\max}^{1\times 2^n - 1} \quad \text{and} \quad h_{\max}^{1\times 2^n - 2}.$$

6.4 Conclusions

Following the motivation presented in Section 1.2, we have generalized in this chapter the basic results from Chapter 2 that laid the foundations for the new and innovative work which comprises Chapters 2, 3, 4 and 5, of this dissertation.

Namely, we have extended the results in Theorems 2.1.4–2.1.5, for continuous and piecewise analytic potentials (1.0.13), to the results in Theorems 6.3.1–6.3.2, for absolutely integrable potentials (1.0.14).

In the course of this generalization, we have seen that the regularity of the potential influences:

- the maximum step size, as reported in Section 6.1, where, for instance, the maximum step size is restricted differently for Class I as (6.1.9)–(6.1.10), whereas for Class III as (6.1.14)–(6.1.15), and,
- the convergence rate, as exposed in Section 6.3, most notably via Corollaries 6.3.1, 6.3.2, 6.3.3, 6.3.4, 6.3.5 and 6.3.6 throughout Classes I, II, III and IV.

Notwithstanding these differences, we have shown that regardless of the regularity of the potential, the extended results remain accompanied by error bounds which have the properties (i) and (ii) from page 25, that welcome a future investigation of this set of ideas in broader settings.

6.5 Proof of Theorem 6.3.1

Similarly to the proof of Theorem 2.1.4, the argument here also follows from the closed-form representations of Fer streamers from Theorem 2.1.3.

6.5.1 Estimating $\exp\left(\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1})\right)\cdots\exp\left(\boldsymbol{D}_{\lambda,0}(a, c_1)\right)$

6.5.1.1 Classes I and II

In this subsubsection it is assumed that q belongs to Class I or to Class II. Recall (2.4.1) and write verbatim:

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) = 2|t-c_k|\sqrt{\frac{\int_{[c_k,t]} q(\xi)d\xi}{|t-c_k|} - \lambda}. \tag{6.5.1}$$

Since q belongs to Classes I or II, note further that (2.4.2)–(2.4.5) hold also for these more general potentials, and can be written, with $q_{\max}$ replaced by $\operatorname{ess\,sup}\{q\}$, concisely as:

$$|\lambda - \operatorname{ess\,sup}\{q\}| \leq h_{\max}^{-2} \Rightarrow |\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)| \leq 2\sqrt{2}, \tag{6.5.2}$$

$$\lambda - \operatorname{ess\,sup}\{q\} \geq h_{\max}^{-2} \Rightarrow \rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right) \in i\Big[2|t-c_k|\sqrt{\lambda - \operatorname{ess\,sup}\{q\}}, +\infty\Big). \tag{6.5.3}$$

Once again, because q lies in Classes I or II, one may also call upon (2.4.6)–(2.4.8), which can be reformulated immediately to yield the estimates in the uniform regime (6.1.12):

$$|\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)|t-c_k|| \leq 2h_{\max}, \tag{6.5.4}$$

$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)|t-c_k|^2\right| \leq \left(2h_{\max}\right)^2, \tag{6.5.5}$$

$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right| \leq 2, \tag{6.5.6}$$

and similarly, with $q_{\max}$ replaced by $\operatorname{ess\,sup}\{q\}$, (2.4.9)–(2.4.11) result immediately in the estimates in the uniform regime (6.1.13):

$$|\varphi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,|t-c_k||\leq(\lambda-\operatorname{ess\,sup}\{q\})^{-\frac{1}{2}}, \tag{6.5.7}$$

$$\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,|t-c_k|^2\right|\leq(\lambda-\operatorname{ess\,sup}\{q\})^{-1}, \tag{6.5.8}$$

$$\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,\rho^2(\boldsymbol{D}_{\lambda,0}(c_k,t))\right|\leq 2. \tag{6.5.9}$$

If q belongs to Class I, observe that assumptions (6.1.5), (6.1.6), (6.1.7) and (6.1.8) and Hölder's inequality imply that

$$\begin{aligned}
&\left|q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right|=\\
&=\left|q_k(t)-\frac{\int_{[c_k,t]}q_k(\xi)d\xi}{|t-c_k|}\right|\ \text{a.e. } t\in[c_k,c_{k+1}]\\
&=\left|\left(q_k(c_k)+\int_{[c_k,t]}q_k'(\xi_2)d\xi_2\right)-\frac{\int_{[c_k,t]}\left(q_k(c_k)+\int_{[c_k,\xi]}q_k'(\xi_2)d\xi_2\right)d\xi}{|t-c_k|}\right|\\
&\leq\int_{[c_k,t]}|q_k'(\xi_2)|d\xi_2+\frac{\int_{[c_k,t]}\int_{[c_k,\xi]}|q_k'(\xi_2)|d\xi_2d\xi}{|t-c_k|}\\
&\leq|t-c_k|^{\frac{\gamma-1}{\gamma}}\|q_k'\|_{\mathrm{L}^\gamma([c_k,c_{k+1}],\mathbb{R})}+\frac{\int_{[c_k,t]}|\xi-c_k|^{\frac{\gamma-1}{\gamma}}\|q_k'\|_{\mathrm{L}^\gamma([c_k,c_{k+1}],\mathbb{R})}d\xi}{|t-c_k|}\\
&=\frac{3\gamma-1}{2\gamma-1}\|q_k'\|_{\mathrm{L}^\gamma([c_k,c_{k+1}],\mathbb{R})}|t-c_k|^{\frac{\gamma-1}{\gamma}}\\
&=\frac{3\gamma-1}{2\gamma-1}\|q'\|_{\mathrm{L}^\gamma([c_k,c_{k+1}],\mathbb{R})}|t-c_k|^{\frac{\gamma-1}{\gamma}}
\end{aligned}$$

and result in

$$\begin{aligned}
&\int_{[c_k,t]}\left|q(\xi)-\frac{\int_{[c_k,\xi]}q(\xi_2)d\xi_2}{|\xi-c_k|}\right|d\xi\leq\\
&\leq\frac{(3\gamma-1)\gamma}{(2\gamma-1)^2}\|q'\|_{\mathrm{L}^\gamma([c_k,c_{k+1}],\mathbb{R})}h_{\max}^{\frac{2\gamma-1}{\gamma}}\\
&\leq\begin{cases}\frac{3}{4}\|q'\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}h_{\max}^2 & \Leftarrow\gamma=\infty,\\ \frac{(3\gamma-1)\gamma}{(2\gamma-1)^2}\|q'\|_{\mathrm{L}^\gamma([a,b],\mathbb{R})}o\left(h_{\max}^{\frac{2\gamma-1}{\gamma}}\right) & \Leftarrow\gamma\in(1,+\infty),\\ 2\|q'\|_{\mathrm{L}^1([a,b],\mathbb{R})}o\left(h_{\max}\right) & \Leftarrow\gamma=1.\end{cases}
\end{aligned}\tag{6.5.10}$$

If q belongs to Class II, observe that Hölder's inequality yields

$$\begin{aligned}\int_{[c_k,t]}\left|q(\xi)-\frac{\int_{[c_k,\xi]}q(\xi_2)d\xi_2}{|\xi-c_k|}\right|d\xi &\leq 2\|q\|_{\mathrm{L}^\infty([c_k,c_{k+1}],\mathbb{R})}h_{\max}\\ &\leq 2\|q\|_{\mathrm{L}^\infty([a,b],\mathbb{R})}h_{\max}. \qquad (6.5.11)\end{aligned}$$

Finally, we are in a position to estimate

$$\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)\cdots\exp\left(\boldsymbol{D}_{\lambda,0}(a,c_1)\right).$$

To this end, we require a different approach for each of the two uniform regimes (6.1.12) and (6.1.13). Firstly, in the uniform regime (6.1.12), we have

$$\begin{aligned}&e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}=\\ &=\cosh\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}\begin{bmatrix}1&0\\0&1\end{bmatrix}+\\ &\quad+\frac{\sinh\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}}{\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}}\begin{bmatrix}0&c_{k+1}-c_k\\(c_{k+1}-c_k)^{-1}\left(\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}\right)^2&0\end{bmatrix}\\ &=\mathcal{O}(1)\begin{bmatrix}1&0\\0&1\end{bmatrix}+\mathcal{O}(1)\begin{bmatrix}0&\mathcal{O}(1)(2h_{\max})\\\mathcal{O}(1)(2h_{\max})^{-1}&0\end{bmatrix}\end{aligned}$$

where we have called upon assumptions (6.1.10) and (6.1.11) as well as (6.5.2). Secondly, in the uniform regime (6.1.13), we have

$$\begin{aligned}&e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})}=\\ &=\cos\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2i}\begin{bmatrix}1&0\\0&1\end{bmatrix}+\\ &\quad+\sin\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2i}\begin{bmatrix}0&\frac{c_{k+1}-c_k}{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\\-\frac{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{c_{k+1}-c_k}&0\end{bmatrix}\\ &=\mathcal{O}(1)\begin{bmatrix}1&0\\0&1\end{bmatrix}+\mathcal{O}(1)\begin{bmatrix}0&\mathcal{O}(1)(\lambda-\operatorname{ess\,sup}\{q\})^{-\frac{1}{2}}\\\mathcal{O}(1)\left((\lambda-\operatorname{ess\,sup}\{q\})^{-\frac{1}{2}}\right)^{-1}&0\end{bmatrix}\end{aligned}$$

where we have taken advantage of (6.5.3) and of the fact that assumption (6.1.9) ensures that

$$\begin{aligned}
\left|\frac{c_{k+1}-c_k}{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\right| &= \frac{1}{\sqrt{\lambda-\frac{\int_{[c_k,c_{k+1}]}q(\xi)d\xi}{c_{k+1}-c_k}}} \\
&\leq 1\cdot\frac{1}{\sqrt{\lambda-\operatorname{ess\,sup}\{q\}}}, \\
\left|\frac{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{c_{k+1}-c_k}\right| &= \sqrt{\lambda-\frac{\int_{[c_k,c_{k+1}]}q(\xi)d\xi}{c_{k+1}-c_k}} \\
&\leq \sqrt{\frac{\lambda-\operatorname{ess\,inf}\{q\}}{\lambda-\operatorname{ess\,sup}\{q\}}}\cdot\sqrt{\lambda-\operatorname{ess\,sup}\{q\}} \\
&\leq \sqrt{1+h_{\max}^2(\operatorname{ess\,sup}\{q\}-\operatorname{ess\,inf}\{q\})}\cdot\sqrt{\lambda-\operatorname{ess\,sup}\{q\}} \\
&\leq \sqrt{2}\cdot\sqrt{\lambda-\operatorname{ess\,sup}\{q\}}.
\end{aligned}$$

The result now follows from Definition 6.3.1.

6.5.1.2 Classes III and IV

In this subsubsection it is assumed that q belongs to either Class III or IV. The treatment follows that of the previous subsection, but presents new subtleties which require additional care. Rewrite (6.5.1) as

$$\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)=2|t-c_k|^{\frac{2p-1}{2p}}\sqrt{|t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi-|t-c_k|^{\frac{1}{p}}\lambda}$$

and observe that assumptions (6.1.14)–(6.1.15) and Hölder's inequality yield

$$\left||t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi\right|\leq\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})} \tag{6.5.12}$$

and

$$|\lambda|\leq h_{\max}^{-2}\left(1-h_{\max}^{\frac{2p-1}{p}}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)\Rightarrow|\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)|\leq 2, \tag{6.5.13}$$

$$\lambda\geq h_{\max}^{-2}\left(1-h_{\max}^{\frac{2p-1}{p}}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)\Rightarrow\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\in[0,2]\cup i\mathbb{R}_0^+. \tag{6.5.14}$$

Like before, (6.5.13), Definition 2.1.4 and Remark 2.1.4, lead to the following estimates in the uniform regime (6.1.17)

$$|\varphi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))|t-c_k||\leq 2h_{\max}, \tag{6.5.15}$$

$$\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))|t-c_k|^2\right|\leq (2h_{\max})^2, \tag{6.5.16}$$

$$\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\rho^2(\boldsymbol{D}_{\lambda,0}(c_k,t))\right|\leq 2. \tag{6.5.17}$$

The new subtlety appears in the uniform regime (6.1.18). Unlike before, (6.5.14) does not lead to 'good' estimates. A possible workaround is to partition

$$[c_k,c_{k+1}]=\left[c_k,c_k+\lambda^{-\frac{1}{2}}\right]\cup\left[c_k+\lambda^{-\frac{1}{2}},c_{k+1}\right].$$

If $t\in\left[c_k,c_k+\lambda^{-\frac{1}{2}}\right]$, then it is clear that (6.5.14) results in

$$|\varphi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))|t-c_k||\leq 2\lambda^{-\frac{1}{2}}, \tag{6.5.18}$$

$$\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))|t-c_k|^2\right|\leq \left(2\lambda^{-\frac{1}{2}}\right)^2, \tag{6.5.19}$$

$$\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\rho^2(\boldsymbol{D}_{\lambda,0}(c_k,t))\right|\leq 2. \tag{6.5.20}$$

If $t\in\left[c_k+\lambda^{-\frac{1}{2}},c_{k+1}\right]$, then it follows from assumption (6.1.14), (6.5.12), (6.5.14) and the inequalities

$$\begin{aligned}
&|t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi-|t-c_k|^{\frac{1}{p}}\lambda\leq\\
&\leq\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}-\lambda^{\frac{2p-1}{2p}}\\
&\leq -h_{\max}^{-\frac{2p-1}{p}}\left(\left(1-h_{\max}^{\frac{2p-1}{p}}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)^{\frac{2p-1}{2p}}-h_{\max}^{\frac{2p-1}{p}}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)\\
&\leq -\frac{h_{\max}^{-\frac{2p-1}{p}}}{2}\\
&<0
\end{aligned}$$

and

$$\left|\frac{|t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|^{\frac{1}{p}}\lambda}\right|\leq\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\lambda^{-\frac{2p-1}{2p}}\leq\frac{h_{\max}^{\frac{2p-1}{p}}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}}{\left(1-h_{\max}^{\frac{2p-1}{p}}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right)^{\frac{2p-1}{2p}}}\leq\frac{1}{3}$$

that

$$\begin{aligned}
|\varphi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,|t-c_k|| &= \left|\frac{\varphi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,\rho(\boldsymbol{D}_{\lambda,0}(c_k,t))}{2}\,\frac{2\,|t-c_k|}{\rho(\boldsymbol{D}_{\lambda,0}(c_k,t))}\right| \\
&\leq \lambda^{-\frac{1}{2}}\left(\frac{|t-c_k|^{\frac{1}{p}}\lambda}{|t-c_k|^{\frac{1}{p}}\lambda - |t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi}\right)^{\frac{1}{2}} \\
&= \lambda^{-\frac{1}{2}}\left(1-\frac{|t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|^{\frac{1}{p}}\lambda}\right)^{-\frac{1}{2}} \\
&\leq 2\lambda^{-\frac{1}{2}},
\end{aligned}$$

$$\begin{aligned}
&\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,|t-c_k|^2\right| = \\
&= \left|\frac{\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,\rho^2(\boldsymbol{D}_{\lambda,0}(c_k,t))}{4}\,\frac{4\,|t-c_k|^2}{\rho^2(\boldsymbol{D}_{\lambda,0}(c_k,t))}\right| \\
&\leq \lambda^{-1}\frac{|t-c_k|^{\frac{1}{p}}\lambda}{|t-c_k|^{\frac{1}{p}}\lambda - |t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi} \\
&= \lambda^{-1}\left(1-\frac{|t-c_k|^{\frac{1-p}{p}}\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|^{\frac{1}{p}}\lambda}\right)^{-1} \\
&\leq \left(2\lambda^{-\frac{1}{2}}\right)^2
\end{aligned}$$

and

$$\left|\phi(\rho(\boldsymbol{D}_{\lambda,0}(c_k,t)))\,\rho^2(\boldsymbol{D}_{\lambda,0}(c_k,t))\right| \leq 2.$$

If q belongs to Class III, then Hölder's inequality yields

$$\begin{aligned}
\int_{[c_k,t]}\left|q(\xi)-\frac{\int_{[c_k,\xi]}q(\xi_2)d\xi_2}{|\xi-c_k|}\right|d\xi &\leq \frac{2p-1}{p-1}\|q\|_{\mathrm{L}^p([c_k,c_{k+1}],\mathbb{R})}h_{\max}^{\frac{p-1}{p}} \\
&\leq \frac{2p-1}{p-1}\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}o\left(h_{\max}^{\frac{p-1}{p}}\right). \qquad (6.5.21)
\end{aligned}$$

If q belongs to Class IV and c_k is a Lebesgue point of q, then Lebesgue's fundamental theorem of calculus ensures that the mapping

$$\xi\in[c_k,t]\to\int_{[c_k,\xi]}|q(\xi_2)|d\xi_2\in\mathbb{R}_0^+$$

is continuous and Lebesgue's differentiation theorem ensures that

$$\exists \lim_{\xi \to c_k^+} \frac{\int_{[c_k,\xi]} |q(\xi_2)| d\xi_2}{|\xi - c_k|} < +\infty.$$

Hence,

$$\xi \in [c_k, t] \to \frac{\int_{[c_k,\xi]} |q(\xi_2)| d\xi_2}{|\xi - c_k|} \in \mathbb{R}_0^+$$

is continuous (with removable singularity) and

$$\begin{aligned}
&\int_{[c_k,t]} \left| q(\xi) - \frac{\int_{[c_k,\xi]} q(\xi_2) d\xi_2}{|\xi - c_k|} \right| d\xi \leq \\
&\leq \|q\|_{\mathrm{L}^1([a,b],\mathbb{R})} \left(\frac{\int_{[c_k,c_{k+1}]} |q(\xi)| d\xi}{\|q\|_{\mathrm{L}^1([a,b],\mathbb{R})}} + \frac{\int_{[c_k,t]} \frac{\int_{[c_k,\xi]} |q(\xi_2)| d\xi_2}{|\xi - c_k|} d\xi}{\|q\|_{\mathrm{L}^1([a,b],\mathbb{R})}} \right) \\
&\leq \|q\|_{\mathrm{L}^1([a,b],\mathbb{R})} \left(o(1) + \mathcal{O}(h_{\max}) \right).
\end{aligned} \tag{6.5.22}$$

Finally, we have the capacity to estimate

$$\exp\left(\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1})\right) \cdots \exp\left(\boldsymbol{D}_{\lambda,0}(a, c_1)\right).$$

To this end we require a different way of dealing with each of the two uniform regimes (6.1.17) and (6.1.18). Firstly, in the uniform regime (6.1.17), we have, like before,

$$\begin{aligned}
&e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} = \\
&= \cosh \frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k, c_{k+1})\right)}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \\
&\quad + \frac{\sinh \frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}}{\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2}} \begin{bmatrix} 0 & c_{k+1} - c_k \\ (c_{k+1} - c_k)^{-1} \left(\frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2} \right)^2 & 0 \end{bmatrix} \\
&= \mathcal{O}(1) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \mathcal{O}(1) \begin{bmatrix} 0 & \mathcal{O}(1)(2h_{\max}) \\ \mathcal{O}(1)(2h_{\max})^{-1} & 0 \end{bmatrix}
\end{aligned}$$

where we have called upon assumptions (6.1.15)–(6.1.16) and (6.5.13). Secondly, in the uniform regime (6.1.18), we have, unlike before,

$$
\begin{aligned}
& e^{\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})} = \\
& = \cos \frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2i} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \\
& \quad + \sin \frac{\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{2i} \begin{bmatrix} 0 & \frac{c_{k+1}-c_k}{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)} \\ -\frac{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{c_{k+1}-c_k} & 0 \end{bmatrix} \\
& = \mathcal{O}(1) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \mathcal{O}(1) \begin{bmatrix} 0 & \mathcal{O}(1)\left(2\lambda^{-\frac{1}{2}}\right) \\ \mathcal{O}(1)\left(2\lambda^{-\frac{1}{2}}\right)^{-1} & 0 \end{bmatrix}
\end{aligned}
$$

where we have capitalized upon (6.5.14) as well as the fact that assumption (6.1.14), assumption (6.1.16) and (6.5.12) ensure that

$$
\begin{aligned}
& (c_{k+1}-c_k)^{\frac{1-p}{p}} \int_{[c_k,c_{k+1}]} q(\xi)d\xi - (c_{k+1}-c_k)^{\frac{1}{p}}\lambda \leq \\
& \leq \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})} - \left(\frac{h_{\min}}{h_{\max}}\right)^{\frac{1}{p}} h_{\max}^{\frac{1}{p}}\lambda \\
& \leq -h_{\max}^{-\frac{2p-1}{p}} \left(\frac{1}{2} - \frac{3}{2} h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}\right) \\
& \leq -\frac{h_{\max}^{-\frac{2p-1}{p}}}{8} \\
& < 0,
\end{aligned}
$$

$$
\begin{aligned}
\left| \frac{(c_{k+1}-c_k)^{\frac{1-p}{p}} \int_{[c_k,c_{k+1}]} q(\xi)d\xi}{(c_{k+1}-c_k)^{\frac{1}{p}}\lambda} \right| &\leq \left(\frac{h_{\max}}{h_{\min}}\right)^{\frac{1}{p}} \frac{\|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}}{h_{\max}^{\frac{1}{p}}\lambda} \\
&\leq 2 \frac{h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}}{1 - h_{\max}^{\frac{2p-1}{p}} \|q\|_{\mathrm{L}^p([a,b],\mathbb{R})}} \\
&\leq \frac{2}{3}
\end{aligned}
$$

and

$$\left|\frac{c_{k+1}-c_k}{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}\right| = \lambda^{-\frac{1}{2}}\left(1-\frac{\left(c_{k+1}-c_k\right)^{\frac{1-p}{p}}\int_{[c_k,c_{k+1}]}q(\xi)d\xi}{\lambda\left(c_{k+1}-c_k\right)^{\frac{1}{p}}}\right)^{-\frac{1}{2}}$$
$$\leq \frac{\sqrt{3}}{2}\cdot\left(2\lambda^{-\frac{1}{2}}\right),$$
$$\left|\frac{(2i)^{-1}\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)}{c_{k+1}-c_k}\right| = \lambda^{\frac{1}{2}}\left(1-\frac{\left(c_{k+1}-c_k\right)^{\frac{1-p}{p}}\int_{[c_k,c_{k+1}]}q(\xi)d\xi}{\lambda\left(c_{k+1}-c_k\right)^{\frac{1}{p}}}\right)^{\frac{1}{2}}$$
$$\leq \frac{2\sqrt{5}}{\sqrt{3}}\cdot\left(2\lambda^{-\frac{1}{2}}\right)^{-1}.$$

The result now follows from Definition 6.3.1.

6.5.2 Estimating $\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)$ and $\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)$

Contrary to the previous subsection, it is now possible and convenient to cover every class and uniform regime simultaneously. To this end, recall Definition 6.3.1 and rewrite (6.5.4)–(6.5.6), (6.5.7)–(6.5.9), (6.5.15)–(6.5.17) and (6.5.18)–(6.5.20) as

$$\left|\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)|t-c_k|\right| \leq \epsilon_1, \tag{6.5.23}$$
$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)|t-c_k|^2\right| \leq \epsilon_1^2, \tag{6.5.24}$$
$$\left|\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right| \leq 2, \tag{6.5.25}$$

and (6.5.10), (6.5.11), (6.5.21) and (6.5.22) as

$$\int_{[c_k,t]}\left|q(\xi)-\frac{\int_{[c_k,\xi]}q(\xi_2)d\xi_2}{|\xi-c_k|}\right|d\xi \leq \epsilon_2. \tag{6.5.26}$$

Note that (6.5.23)–(6.5.25), in turn, imply that

$$\begin{aligned}
&\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right)= \\
&=\begin{bmatrix}\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)|t-c_k|\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\\ 0\\ 0\end{bmatrix} \\
&=\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ 0\\ 0\end{bmatrix}
\end{aligned}$$

and

$$\begin{aligned}
&\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right)= \\
&=\begin{bmatrix}0\\ -2\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)|t-c_k|^2\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\\ \frac{1}{2}\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\rho^2\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\end{bmatrix} \\
&=\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\begin{bmatrix}0\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}
\end{aligned}$$

which, according to Theorem 2.1.3, lead to

$$\begin{aligned}
\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)&=\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right)+ \\
&\quad+\phi\left(\rho\left(\boldsymbol{D}_{\lambda,0}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,0}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,0}(c_k,t)\right) \\
&=\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}
\end{aligned}$$

and (c.f., (6.5.26))

$$\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)=\int_{[c_k,t]}\boldsymbol{B}_{\lambda,1}(c_k,\xi)d\xi=\epsilon_2\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}.$$

6.5.3 Estimating $\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right)$ and $\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)$ for $l\geq 2$

Our estimate follows by induction. The induction claim is that

$$\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right)=\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\epsilon_2^{2^{l-1}-1}\epsilon_1^{2^{l-1}-1}\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix},$$

$$\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)=\epsilon_2^{2^{l-1}}\epsilon_1^{2^{l-1}-1}\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}.$$

6.5.3.1 First step: $l=2$

Given Definition 2.1.4 and the uniform estimates for $\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)$ in the previous subsection, it is now clear that

$$\begin{aligned}\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right)&=-\frac{1}{2}+\epsilon_2^2\mathcal{O}\left(\epsilon_1^2\right),\\ \phi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right)&=\frac{1}{3}+\epsilon_2^2\mathcal{O}\left(\epsilon_1^2\right),\end{aligned}$$

and, according to Theorem 2.1.3, that

$$\begin{aligned}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,2}(c_k,t)\right)&=\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)+\\ &\quad+\phi\left(\rho\left(\boldsymbol{D}_{\lambda,1}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,1}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,1}(c_k,t)\right)\\ &=\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\epsilon_2\epsilon_1\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}\end{aligned}$$

and (c.f., (6.5.26))

$$\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,2}(c_k,t)\right)=\int_{[c_k,t]}\boldsymbol{B}_{\lambda,2}(c_k,\xi)d\xi=\epsilon_2^2\epsilon_1\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}.$$

6.5.3.2 Induction step: $l\Rightarrow l+1$

Given the induction claim, it is now clear that

$$\begin{aligned}\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right)&=-\frac{1}{2}+\epsilon_2^{2^l}\mathcal{O}\left(\epsilon_1^{2^l}\right),\\ \phi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right)&=\frac{1}{3}+\epsilon_2^{2^l}\mathcal{O}\left(\epsilon_1^{2^l}\right),\end{aligned}$$

and, according to Theorem 2.1.3, that

$$\begin{aligned}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l+1}(c_k,t)\right)&=\varphi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right)\mathscr{C}_{\boldsymbol{D}_{\lambda,l}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right)+\\&\quad+\phi\left(\rho\left(\boldsymbol{D}_{\lambda,l}(c_k,t)\right)\right)\mathscr{C}^2_{\boldsymbol{D}_{\lambda,l}(c_k,t)}\boldsymbol{\pi}\left(\boldsymbol{B}_{\lambda,l}(c_k,t)\right)\\&=\left(q(t)-\frac{\int_{[c_k,t]}q(\xi)d\xi}{|t-c_k|}\right)\epsilon_2^{2^l-1}\epsilon_1^{2^l-1}\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}\end{aligned}$$

and (c.f., (6.5.26))

$$\boldsymbol{\pi}\left(\boldsymbol{D}_{\lambda,l+1}(c_k,t)\right)=\int_{[c_k,t]}\boldsymbol{B}_{\lambda,l+1}(c_k,\xi)d\xi=\epsilon_2^{2^l}\epsilon_1^{2^l-1}\begin{bmatrix}\mathcal{O}\left(\epsilon_1\right)\\ \mathcal{O}\left(\epsilon_1^2\right)\\ \mathcal{O}\left(1\right)\end{bmatrix}.$$

6.6 Proof of Theorem 6.3.2

As in the proof of Theorem 2.1.5, the main obstacle in estimating the local and global errors is the fact that the lower-left entry of $\exp\left(\boldsymbol{D}_{\lambda,0}(c_k,c_{k+1})\right)$ is very large. This is circumvented by calling upon three Baker–Campbell–Hausdorff (BCH) type formulas (2.5.1), (2.5.2) and (2.5.3). Firstly, the local error is estimated by calling upon Definition 2.1.6, the aforementioned BCH type formulas and Theorem 6.3.1. Secondly, the global error is estimated by invoking Definition 2.1.6, the BCH formulas (2.5.1)–(2.5.3), Theorem 6.3.1 as well as assumption (6.1.11) (when dealing with Classes I or II) or assumption (6.1.16)

(when dealing with Classes III or IV). This is done by observing that the global error obeys a certain recurrence relation.

Bibliography

Amrein, W. O., Hinz, A. M. and Pearson, D. P., eds (2005), *Sturm–Liouville Theory: Past and Present*, Birkhäuser Verlag Basel. (Cited on page 2).

Anderssen, R. and de Hoog, F. (1984), 'On the correction of finite difference eigenvalue approximations for Sturm–Liouville problems with general boundary conditions', *BIT Numerical Mathematics* **24**(4), 401–412. (Cited on page 73).

Brent, R. (2002), *Algorithms for minimization without derivatives*, reprint edn, Dover, New York. (Cited on page 100).

Degani, I. (2004), RCMS - Right Correction Magnus Schemes for Oscillatory ODEs, and Cubature Formulae and Commuting Extensions, PhD thesis, Weizmann Institute of Science, Department of Mathematics. (Cited on pages 7, 10, 11, 12, 13, 14, and 15).

Degani, I. and Schiff, J. (2006), 'RCMS: Right Correction Magnus Series approach for oscillatory ODEs', *Journal of Computational and Applied Mathematics* **193**(2), 413–436. (Cited on pages 7, 10, 11, 12, 13, 14, 15, 18, 19, 20, 21, and 77).

Engø, K., Marthinsen, A. and Munthe–Kaas, H. Z. (1999), DiffMan — an object oriented MATLAB toolbox for solving differential equations on manifolds, Technical report, Department of Computer Science, University of Bergen, Norway. Available at: (http://www.diffman.no/). (Cited on page 92).

Evans, M., Coffey, W. and Pryce, J. D. (1979), 'The effect of dipole-dipole interaction on zero-thz frequency polarisation', *Chemical Physics Letters* **63**(1), 133–138. (Cited on page 74).

Fer, F. (1958), 'Résolution del l'equation matricielle $\dot{U} = pU$ par produit infini d'exponentielles matricielles', *Bulletin de la Classe des Sciences Académie Royale de Belgique* **44**, 818–829. (Cited on pages 26 and 28).

Iserles, A. (1984), 'Solving linear ordinary differential equations by exponentials of iterated commutators', *Numerische Mathematik* **45**(2), 183–199. (Cited on pages 26 and 28).

Iserles, A. (2004*a*), 'On the method of Neumann series for highly oscillatory equations', *BIT Numerical Mathematics* **44**(3), 473–488. (Cited on pages 7, 10, 12, 13, 14, and 15).

Iserles, A. (2004*b*), 'On the numerical quadrature of highly-oscillating integrals I: Fourier transforms', *IMA Journal of Numerical Analysis* **24**(3), 365–391. (Cited on pages 7, 10, 11, 12, 13, 14, 15, and 53).

Iserles, A., Munthe–Kaas, H. Z., Nørsett, S. and Zanna, A. (2000), 'Lie-group methods', *Acta Numerica* **9**, 215–365. (Cited on pages 14, 26, 27, 28, and 95).

Iserles, A. and Nørsett, S. (1999*a*), 'On the solution of linear differential equations in Lie groups', *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* **357**(1754), 983–1019. (Cited on page 53).

Iserles, A. and Nørsett, S. (1999*b*), 'On the solution of linear differential equations on Lie-groups', *Philosophical Transactions of the Royal Society A* **357**, 983–1019. (Cited on page 13).

Iserles, A. and Nørsett, S. (2005), 'Efficient quadrature of highly oscillatory integrals using derivatives', *Proceedings of the Royal Society A* **461**, 1383–1399. (Cited on pages 54 and 67).

Iserles, A. and Nørsett, S. (2006), 'Quadrature methods for multivariate highly oscillatory integrals using derivatives', *Mathematics of Computation* **75**(255), 1233–1258. (Cited on pages 12, 54, and 67).

Ixaru, L. G. (2000), 'CP methods for the Schrödinger equation', *Journal of Computational and Applied Mathematics* **125**(1-2), 347–357. (Cited on pages 7, 8, 9, 10, 12, 13, 14, and 15).

Ixaru, L. G., De Meyer, H. and Berghe, G. V. (1997), 'CP methods for the Schrödinger equation revisited', *Journal of Computational and Applied Mathematics* **88**(2), 289–314. (Cited on pages 6, 7, 8, 9, 10, 12, 13, 14, 15, and 101).

Ixaru, L. G., De Meyer, H. and Berghe, G. V. (1999), 'SLCPM12 — A program for solving regular Sturm–Liouville problems', *Computer Physics Communications* **118**(2-3), 259–277. (Cited on pages 4, 7, 8, 9, 10, 12, 13, 14, 15, 99, and 100).

Ledoux, V. and Daele, M. V. (2010), 'Solution of Sturm–Liouville problems using modified Neumann schemes', *SIAM Journal on Scientific Computing* **32**(2), 563–584. (Cited on pages 7, 10, 12, 13, 14, and 15).

Ledoux, V., Daele, M. V. and Berghe, G. V. (2004), 'CP methods of higher order for Sturm–Liouville and Schrödinger equations', *Computer Physics Communications* **162**(3), 151–165. (Cited on pages 7, 8, 9, 10, 12, 13, 14, and 15).

Ledoux, V., Daele, M. V. and Berghe, G. V. (2005), 'MATSLISE: A MATLAB package for the numerical solution of Sturm–Liouville and Schrödinger equations', *ACM Transactions on Mathematical Software* **31**(4), 532–554. (Cited on pages 7, 8, 9, 10, 12, 13, 14, 15, 74, and 103).

Ledoux, V., Daele, M. V. and Berghe, G. V. (2010), 'Efficient numerical solution of the 1D Schrödinger eigenvalue problem using Magnus integrators', *IMA Journal of Numerical Analysis* **30**, 751–776. (Cited on pages 7, 10, 11, 12, 13, 14, 15, 18, 19, 20, 21, 77, and 101).

Levin, D. (1996), 'Fast integration of rapidly oscillatory functions', *Journal of Computational and Applied Mathematics* **67**(1), 95–101. (Cited on pages 54 and 67).

Marletta, M. and Pryce, J. D. (1992), 'Automatic solution of Sturm–Liouville problems using the Pruess method', *Journal of Computational and Applied Mathematics* **39**(1), 57–78. (Cited on pages 6, 7, 14, and 15).

Moan, P. C. (1998), Efficient approximation of Sturm–Liouville problems using Lie-group methods, Technical report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom. (Cited on pages 7, 13, and 14).

Munthe–Kaas, H. and Owren, B. (1999), 'Computations in a free lie algebra', *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **357**(1754), 957–981. (Cited on pages 92 and 95).

Olver, F. W. J., Lozier, D. W., Boisvert, R. F. and Clark, C. W., eds (2010), *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY. (Cited on page 68).

Paine, J. and de Hoog, F. (1980), 'Uniform estimation of the eigenvalues of Sturm–Liouville problems', *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics* **21**(3), 365–383. (Cited on pages 6, 7, 8, 14, and 15).

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2007), *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3 edn, Cambridge University Press, New York, NY, USA. (Cited on page 94).

Pruess, S. (1973), 'Estimating the eigenvalues of Sturm–Liouville problems by approximating the differential equation', *SIAM Journal on Numerical Analysis* **10**(1), 55–68. (Cited on pages 6, 7, 8, 14, and 15).

Pruess, S. and Fulton, C. T. (1993), 'Mathematical software for Sturm–Liouville problems', *ACM Transactions on Mathematical Software* **19**(3), 360–376. (Cited on pages 4, 6, 7, 14, 15, 99, and 100).

Pryce, J. D. (1993), *Numerical Solution of Sturm–Liouville Problems*, Oxford University Press. (Cited on pages 1, 2, 74, and 99).

Ramos, A. G. C. P. (2014), Numerical solution of Sturm–Liouville problems via Fer streamers: absolutely integrable potentials and self-adjoint boundary conditions, Technical report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom. Submitted. (Cited on pages iii, 7, 21, 22, 23, and 107).

Ramos, A. G. C. P. (2015*a*), Uniform and high-order discretization schemes for Sturm–Liouville problems via Fer streamers, Technical report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom. Submitted. (Cited on pages iii, 6, 7, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 27, 29, 46, 52, 53, 59, 60, 61, 62, 63, 64, 65, 71, 72, and 73).

Ramos, A. G. C. P. (2015*b*), Uniform and high-order practical implementation of Sturm–Liouville problems via Fer streamers, Technical report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom. Submitted. (Cited on pages iii, 6, 7, 12, 15, 22, 23, 83, 87, and 130).

Ramos, A. G. C. P. (2015*c*), 'Uniform and high-order MATLAB software for Sturm–Liouville problems via Fer streamers'. Provided as supplementary material to (Ramos, 2015*b*). Available at: (http://www.damtp.cam.ac.uk/user/agcpr2/documents/SL_via_Fer_streamers.zip). (Cited on pages iii, 6, 7, 12, 15, 22, 23, 99, and 101).

Ramos, A. G. C. P. and Iserles, A. (2015), 'Numerical solution of Sturm–Liouville problems via Fer streamers', *Numerische Mathematik* **131**(3), 541–565. (Cited on pages iii, 6, 7, 12, 13, 14, 15, 16, 17, 21, 22, 23, 25, 27, 29, 30, 31, 32, 33, and 34).

Zanna, A. (1996), The method of iterated commutators for ordinary differential quations on Lie groups, Technical report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom. (Cited on page 27).

Zanna, A. (1998), On the Numerical Solution of Isospectral Flows, PhD thesis, University of Cambridge, United Kingdom. (Cited on page 28).

Zettl, A. (2005), *Sturm–Liouville Theory*, American Mathematical Society. (Cited on pages 1, 2, 3, 4, and 99).